# By mistake, dropout for all of these was disabled

# Using 4096 neurons

# (weight\_decay, dropout, lr, train\_acc, val\_acc)

# (0.02, 0.6, 0.001, 0.9779, 0.6776)

# (0.025, 0.6, 0.001, 0.7872, 0.6505)

# (0.03, 0.6, 0.001, 0.8064, 0.6620)

# (0.04, 0.6, 0.001, 0.7106, 0.6594)

# (0.05, 0.6, 0.001, 0.7074, 0.6146)

# (0.045, 0.6, 0.001, 0.6857, 0.6255)

# Using 512 neurons

# (0.005, 0.6, 0.001, 0.995, 0.72)

# (0.01, 0.6, 0.001, 0.9969, 0.7042)

# (0.05, 0.6, 0.001, 0.6256, 0.5984)

# (0.02, 0.6, 0.001, 0.92, 0.67)

# (0.03, 0.6, 0.001, 0.8237, 0.6839)

# Using 256 neurons

# (0.01, 0.5, 0.001, 0.9802, 0.7198)

# (0.02, 0.5, 0.001, 0.8936, 0.7182)

# (0.05, 0.5, 0.001, ???) # cancelled

# with dropout enabled

# (weight\_decay, dropout, lr, train\_acc, val\_acc)

# Using 512 neurons (600 batches)

# (0.0001, 0.2, 0.05, 0.1037, 0.0885) -- no learning at all

# Using 1024 neurons (600 batches)

# (0.0001, 0.2, 0.05, inf, inf)

# (0.0001, 0.2, 0.03, 0.9688, 0.7505)

# (0.0001, 0.4, 0.03, 0.1012, 0.0828) -- no learning at all

# (0.0001, 0.3, 0.03, 0.9892, 0.7812) \*\*\*\*\*\*

# (0.0001, 0.35, 0.03, 0.9762, 0.7635)

# (0.0003, 0.3, 0.03, 0.9836, 0.7719)

# (0.0008, 0.3, 0.03, 0.9761, 0.7693)

# (0.001, 0.3, 0.03, 0.9761, 0.7464)

# (0.0008, 0.35, 0.03, ) -- no learning

# (0.003, 0.3, 0.03, 0.9591, 0.7625)

# (0.009, 0.3, 0.03, 0.8398, 0.7369)

# (0.005, 0.3, 0.03, 0.8741, 0.7119)

# (0.007, 0.3, 0.03, 0.8635, 0.6844)

# Using 2048 neurons (600 batches)

# (0.0001, 0.3, 0.03, 0.9977, 0.7755)

# (0.0003, 0.3, 0.03, 0.9983, 0.7745)

# (0.0008, 0.3, 0.03, 0.9875, 0.7740)

# (0.001, 0.3, 0.03, 0.9971, 0.7797)

# (0.005, 0.3, 0.03, 0.9659, 0.7229)

# (0.005, 0.4, 0.03, 0.9659, 0.7229)

# (0.004, 0.6, 0.03, ) -- no learning at all

# (0.00001, 0.6, 0.03, 0.9931, 0.7953)

# (0.0001, 0.6, 0.03, 0.9960, 0.8141) \*\*\*\*\*\*

# (0.001, 0.6, 0.05, 0.9790, 0.8125)

# (0.001, 0.7, 0.05, ...) -- no learning at all

# (0.005, 0.6, 0.05, ) -- no learning

# (0.003, 0.6, 0.05, ) -- no learning

# (0.0008, 0.5, 0.03, 0.9872, 0.7974)

# (0.0008, 0.6, 0.03, 0.9790, 0.8078)

# (0.001, 0.6, 0.03, 0.9857, 0.8135)

# (0.001, 0.6, 0.03, 0.9613, 0.8125)

# Using 2048 neurons (all batches)

# (0.001, 0.6, 0.03, ) -- no learning

# (0.001, 0.6, 0.05, ) -- no learning

# (0.001, 0.6, 0.09, ) -- no learning

# (0.001, 0.6, 0.01, 0.9947, 0.8281) // factor 0.2

# (0.005, 0.6, 0.01, 0.9398, 0.8109) //factor 0.2

# (0.008, 0.6, 0.01, 0.9423, 0.8229) //factor 0.2

# (0.01, 0.6, 0.01, 0.9186, 0.8115) //factor 0.3

# (0.03, 0.6, 0.01, 0.5769, 0.5734) //factor 0.3 -- slow learning

# (0.01, 0.6, 0.01, 0.9202, 0.8099) //factor 0.3

# (0.01, 0.6, 0.01, 0.8212, 0.7812) //factor 0.4; rotation 0.2

# (0.01, 0.6, 0.01, 0.7828, 0.7500) //factor 0.5; rotation 0.2

# (0.01, 0.5, 0.01, 0.8439, 0.7771) //factor 0.5; rotation 0.2

# (0.01, 0.5, 0.01, 0.8698, 0.7589) //factor 0.7; rotation 0.2

# (0.008, 0.6, 0.01, 0.8535, 0.8047) //factor 0.2; rotation 0.2

# (0.008, 0.6, 0.01, 0.8331, 0.7943) //factor 0.5; rotation 0.2

# (0.008, 0.5, 0.008, 0.9043, 0.7802) //factor 0.7; rotation 0.2

# (0.008, 0.6, 0.008, 0.8915, 0.8068) //factor 0.7; rotation 0.2

# (0.008, 0.6, 0.008, 0.9351, 0.8099) //factor 0.7

# (0.01, 0.6, 0.008, 0.9426, 0.7974) //factor 0.7

# (0.03, 0.6, 0.008, 0.8042, 0.7016) //factor 0.7

# (0.05, 0.6, 0.008, ) //factor 0.8 -- no training

# (0.05, 0.6, 0.005, 0.5003, 0.5063) //factor 0.8; slow training

# (0.05, 0.6, 0.007, 0.6478, 0.6406) //factor 0.8; slow learning

# (0.04, 0.6, 0.007, 0.8144, 0.6979) //factor 0.8; slow learning

# (0.005, 0.6, 0.008, 0.9643, 0.8115) // factor 0.7

# (0.01, 0.6, 0.008, 0.9509, 0.7995) // factor 0.7

# 1024 neurons

# (0.01, 0.6, 0.008, 0.9094, 0.7995) // factor 0.7

# (0.03, 0.6, 0.008, 0.7558, 0.6990) // factor 0.7; epochs 20

# (0.02, 0.6, 0.008, 0.8959, 0.7599) // factor 0.7; epochs 30

# (0.025, 0.6, 0.008, 0.8724, 0.7568) // factor 0.7; epochs 30

# (0.028, 0.6, 0.008, 0.8124, 0.7130) // factor 0.8; epochs 30

# (0.001, 0.6, 0.008, 0.9039, 0.8094) // factor 0.7

# (0.005, 0.6, 0.008, 0.9682, 0.8182) // factor 0.7

# (0.01, 0.6, 0.008, 0.9454, 0.8151) // factor 0.7

# 512 neurons

# (0.01, 0.6, 0.008, 0.9252, 0.8255) // factor 0.7

***# (0.01, 0.6, 0.008, 0.9310, 0.8411) // factor 0.5***

***# (0.01, 0.6, 0.008, 0.9389, 0.8365) // factor 0.5***

***# (0.015, 0.6, 0.008,*** ***0.9074, 0.8005) // factor 0.5***

# (0.01, 0.6, 0.008, 0.9174, 0.8250) // factor 0.3

# using 1024 neurons

# (0.01, 0.6, 0.008, 0.9220, 0.8281) // factor 0.5

*# (0.015, 0.6, 0.008, 0.8796, 0.8182) // factor 0.5 epochs 30*

# (0.02, 0.6, 0.008, 0.8156, 0.7641) // factor 0.5 epochs 20

# (0.02, 0.6, 0.008, 0.8850, 0.7922) // factor 0.5 epochs 30

# (0.023, 0.6, 0.008, 0.8218, 0.7781) // factor 0.5

# (0.025, 0.6, 0.008, ) // factor 0.5

# Using 4096 neurons

# (0.01, 0.6, 0.01, 0.9411, 0.7990) //factor 0.3

# (0.02, 0.6, 0.01, 0.8798, 0.7531) //factor 0.3

# (0.01, 0.6, 0.01, 0.8798, 0.7531) //factor 0.3