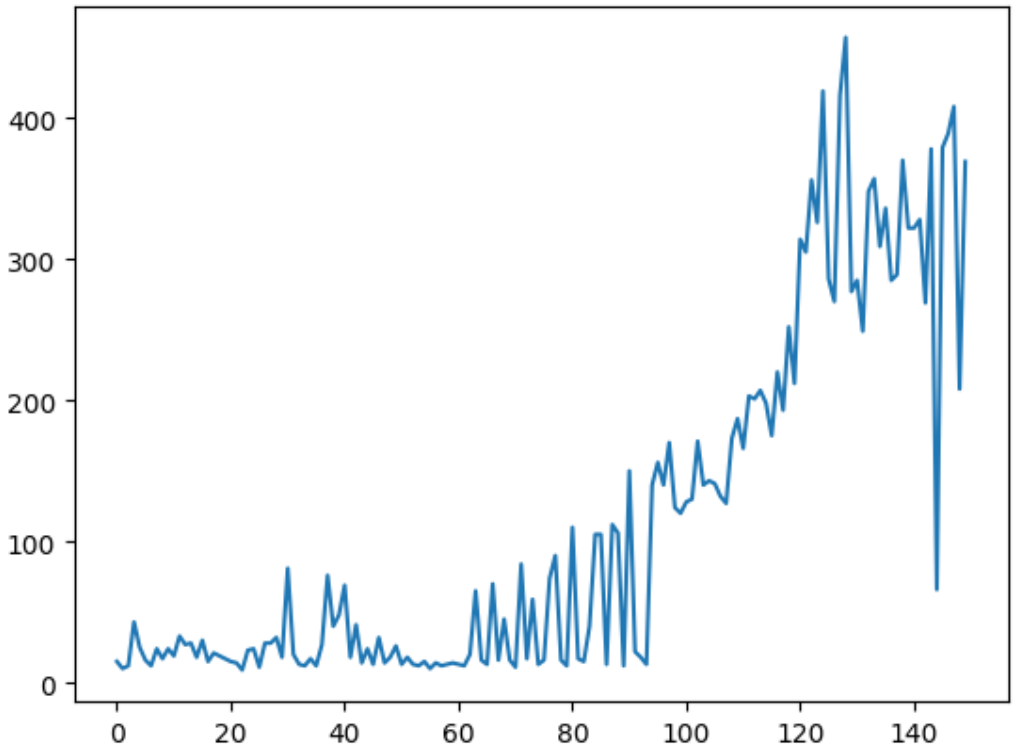
PROBLEM 2:

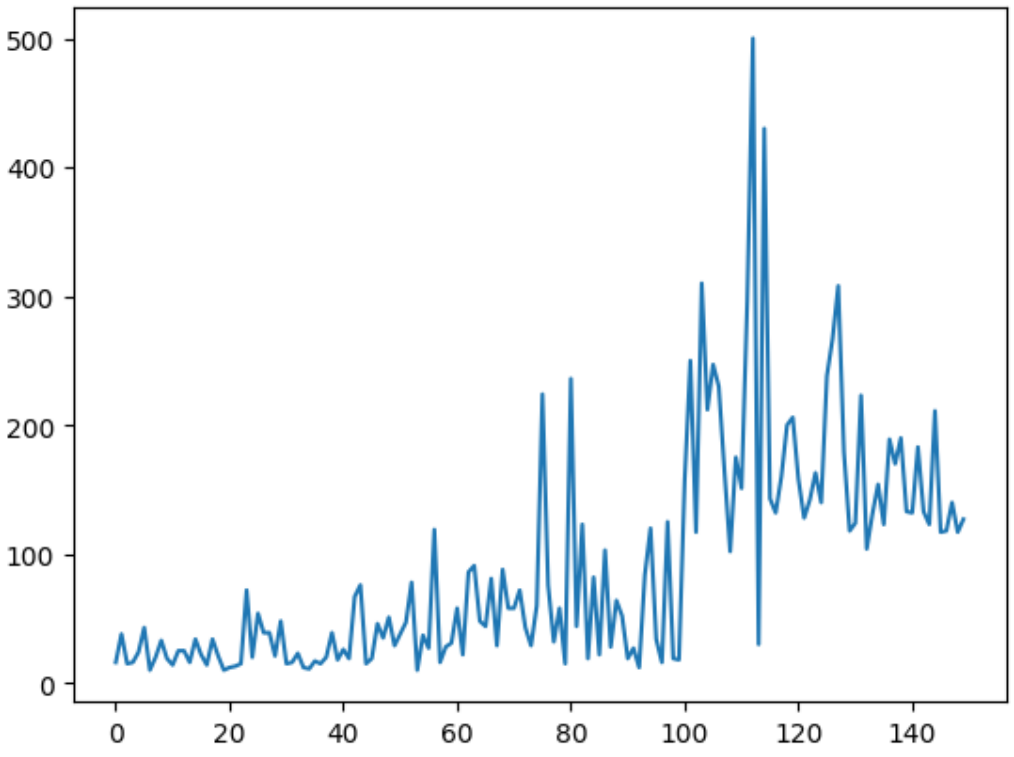
DQN Model:

BATCH_SIZE = 128
GAMMA = 0.999
EPS_START = 0.9
EPS_END = 0.05
EPS_DECAY = 2000
TARGET_UPDATE = 10
MEMORY_CAPACITY = 10000
optimizer = optim.RMS(policy_net.parameters(), lr=0.0025)
loss_fn = nn.SmoothL1Loss()
num_episodes = 150

For the final 10 test environments, the times that occurred are above **380** for all the 10 ten environments.



Double DQN Model:
BATCH_SIZE = 128
GAMMA = 0.999
EPS_START = 0.9
EPS_END = 0.05
EPS_DECAY = 2000
TARGET_UPDATE = 10
MEMORY_CAPACITY = 10000
optimizer = optim.RMSprop(policy_net.parameters(), lr=0.0025)
loss_fn = nn.SmoothL1Loss()
num_episodes = 150



For the final 10 test environments, the times that occurred are above **350** for all the 10 ten environments.