

# Deep Learning

Names: Jeyakumar Sriram (p2214618), Shawn Lim Jun Jie (p2239745)

Class: DAAA/FT/2A/01

Module Code: ST1504

---

## Objective:

Implement a Deep Q-Network (DQN) to control the Gym Pendulum environment, with the goal of training an agent to stabilize the pendulum in an upright position.

## Background:

Reinforcement learning has gained significant attention for solving complex control problems. The Gym Pendulum environment is good for practicing algorithms in continuous control tasks.

## The Gym Pendulum Environment:

The Gym Pendulum environment simulates a simple pendulum with continuous action space. The goal is to train an agent to apply torque to the pendulum, allowing it to stay upright. The state space consists of angular velocity and position, while the action space involves applying torque to the joint.

## Approach:

We will implement the DQN algorithm using PyTorch. The agent will learn to make decisions based on the current state of the pendulum, aiming to maximize cumulative rewards over time. The project will explore how well the DQN performs in mastering the control of the Gym Pendulum.

---

## Project Structure:

1. Imports
  2. Exploring Pendulum Environment
  3. Implementation of DQN with Theory
  4. Model Improvement
    - Using best practices
    - Training Stabilization
    - Enhancing Action Space
    - Enhancing Optimization Strategy
    - Enhancing Model Architecture
  5. Evaluation of Model
  6. Conclusion
- 

## References

- Mnih, V. et al. (2013) 'Playing Atari with Deep Reinforcement Learning', arXiv [cs.LG]. Available at: <http://arxiv.org/abs/1312.5602>.
- GeeksforGeeks. (2021). Bellman Equation. [online] Available at: <https://www.geeksforgeeks.org/bellman-equation/>.
- pytorch.org. (n.d.). Reinforcement Learning (DQN) Tutorial — PyTorch Tutorials 1.8.0 documentation. [online] Available at: [https://pytorch.org/tutorials/intermediate/reinforcement\\_q\\_learning.html](https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html).
- livebook.manning.com. (n.d.). liveBook · Manning. [online] Available at: <https://livebook.manning.com/concept/reinforcement-learning/target-network> [Accessed 27 Jan. 2024].
- Wikipedia. (2020). Circular buffer. [online] Available at: [https://en.wikipedia.org/wiki/Circular\\_buffer](https://en.wikipedia.org/wiki/Circular_buffer).
- baeldung (2020). Epsilon-Greedy Q-learning | Baeldung on Computer Science. [online] www.baeldung.com. Available at: <https://www.baeldung.com/cs/epsilon-greedy-q-learning>.
- Sewak, M. (06 2019) 'Deep Q Network (DQN), Double DQN, and Dueling DQN: A Step Towards General Artificial Intelligence', in, pp. 95–108. doi: 10.1007/978-981-13-8285-7\_8.
- Wikipedia. (2023). Huber loss. [online] Available at: [https://en.wikipedia.org/wiki/Huber\\_loss#:~:text=The%20Huber%20loss%20is%20the](https://en.wikipedia.org/wiki/Huber_loss#:~:text=The%20Huber%20loss%20is%20the) [Accessed 27 Jan. 2024].
- Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236)
- van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double Q-Learning. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16), 2094-2100. Retrieved from <https://arxiv.org/abs/1509.06461>
- Lin, L. J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. Machine Learning, 8(3-4), 293-321. DOI: [10.1007/BF00992699](https://doi.org/10.1007/BF00992699)

## Imports

In [21]:

```
import torch
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torchrl.modules import NoisyLinear

import gymnasium as gym
import math
import sys
import pygame
import numpy as np
import random
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
from collections import namedtuple, deque
from itertools import count
```

In [30]:

```
def showImage(path):
    img = mpimg.imread(path)
    plt.imshow(img)
    plt.axis('off')
    plt.show()

def plot_tuples(tuples_list, title):
    plt.figure()

    for array, name in tuples_list:
        plt.plot(array, label=name)

    plt.xlabel('Episodes')
    plt.ylabel('Score')
    plt.title('Comparison')
    plt.legend(title)
    plt.show()
```

## Exploring the Pendulum Environment

We will be trying to balance the pendulum with its center of gravity above the fixed point



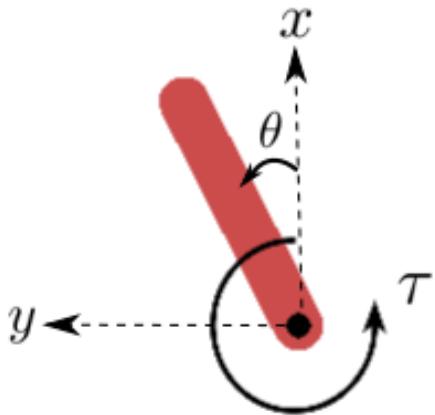
In [2]:

```
env = gym.make("Pendulum-v1", g=9.81)
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
```

## Observation Space

The environment gives us the following information about the pendulum that we can pass into the model.

( $x, y, \theta$ ) where  $x$  and  $y$  refer to the coordinate of the pendulum's end as shown below and  $\theta$  is the angle shown below in radians.



```
In [3]: print(f"Number of Observations: {env.observation_space.shape}")
print(f"Minimum Value for Observations: {env.observation_space.low}")
print(f"Maximum Value for Observations: {env.observation_space.high}")
```

Number of Observations: (3,)  
 Minimum Value for Observations: [-1. -1. -8.]  
 Maximum Value for Observations: [1. 1. 8.]

## Action Space

Our model will generate just one variable T which ranges from -2.0 to 2.0. T is the torque measured in Nm to be applied on the pendulum. A positive T means we are turning it counter-clockwise and vice versa. So our DQN model should produce just one output in this range. Since a DQN doesn't work with Scalar Values, we will discretise this.

```
In [4]: print(f"Shape of Input: {env.action_space.shape}")
print(f"Minimum Value for Input: {env.action_space.low}")
print(f"Maximum Value for Input: {env.action_space.high}")
```

Shape of Input: (1,)  
 Minimum Value for Input: [-2.]  
 Maximum Value for Input: [2.]

## Implementation of DQN (with Theory)

Deep Q-Network (DQN) is a reinforcement learning algorithm that combines deep neural networks with Q-learning, a classic reinforcement learning technique. Developed by researchers at DeepMind, DQN has proven to be effective in solving complex decision-making problems, particularly in domains with high-dimensional state spaces.

## Key Concepts

### Q-Learning

Q-learning is a model-free reinforcement learning algorithm used to learn a policy for making decisions in an environment. The central idea is to approximate the optimal action-value function, denoted as  $Q(s, a)$ , which represents the expected cumulative future rewards when taking action ' $a$ ' in state ' $s$ '.

### Deep Neural Networks

Since Neural Networks are good function approximators, DQN utilizes them to approximate the action-value function. The neural network takes the state of the environment as input and outputs

Q-values for each possible action.

## Experience Replay

To stabilize training, DQN employs experience replay, a technique that stores and randomly samples past experiences from a replay buffer. This helps break the temporal correlations in the data, making the learning process more robust. By using a fixed size circular buffer, we can also ensure only the most recent n values are being learnt upon.

## Loss Function

The loss function used in DQN is typically the mean squared error between the predicted Q-values and the target Q-values. This loss is minimized during training to improve the accuracy of the Q-value predictions.

## MLP

In [6]: # Our Policy Network is a Simple Multi Level Perceptron

```
class BaseMLP(nn.Module):
    def __init__(self, inp, out, lr):
        super(BaseMLP, self).__init__()

        self.layer1 = nn.Linear(inp, 32)
        self.layer2 = nn.Linear(32, 64)
        self.layer3 = nn.Linear(64, 32)
        self.layer4 = nn.Linear(32, out)

        self.lr = lr
        self.optimizer = optim.Adam(self.parameters(), lr=self.lr)

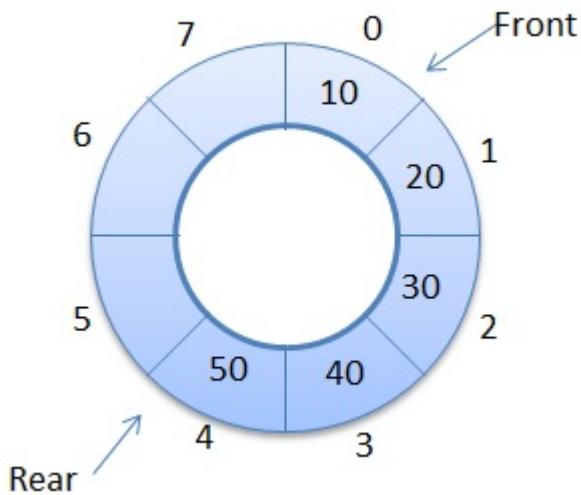
    def forward(self, x):
        x = self.layer1(x)
        x = F.relu(x)
        x = self.layer2(x)
        x = F.relu(x)
        x = self.layer3(x)
        x = F.relu(x)
        return self.layer4(x)
```

## Memory

When we use a DQN we need to have a way for the agent to remember past transitions. Transitions are, in simple english "what was there, what you did, what happened, is it good?". Technically it is a tuple with state, action, next\_state, and reward.

Agent will not use the most recent memory only. Instead we will employ some random sampling to stabilise the training so the model doesn't get caught up with short term patterns.

A good data structure for this will be a circular buffer. This is a queue data struct which follows a FIFO principle. One important thing to note here is that once the buffer is full, subsequent writes will overwrite existing data starting from the oldest. This helps to remove extremely old data.



When we take past memory to train the model, we will also take it in batches. Of course, the samples will be picked randomly.

```
In [7]: class Memory(object):

    def __init__(self, maxlen):
        self.buffer = deque([], maxlen=maxlen)

    def push(self, transition):
        self.buffer.append(transition)

    def sample(self, batch_size):
        mini_batch = random.sample(self.buffer, batch_size)
        states, actions, rewards, next_states = [], [], [], []
        # perform some manipulation to ensure dtypes match
        for transition in mini_batch:
            s, a, r, ns = transition
            states.append(s) # s is a tensor not a regular element
            actions.append([a])
            rewards.append([r])
            next_states.append(ns)

        states = torch.stack(states)
        actions = torch.tensor(actions, dtype=torch.float)
        rewards = torch.tensor(rewards, dtype=torch.float)
        next_states = torch.stack(next_states)
        return states, actions, rewards, next_states

    def __len__(self):
        return len(self.buffer)
```

## DQN Agent

We have a build a very simple DQN with no additional improvements here.

### 1. Initialization

- **Memory Buffer:** A replay memory buffer ( `Memory` ) stores past experiences for experience replay during training.
- **Neural Network:** The Q-function approximator (policy net) takes the state as input and outputs Q-values for each action.

### 2. Action Selection

- The `action` method selects actions based on Q-values predicted by the policy net. The selected action is scaled to match the desired action space.

### 3. Training Optimization

- The `optimize` method updates Q-function parameters using experiences sampled from the replay memory.
  - **Mini-Batch Sampling:** A mini-batch of experiences (state, action, reward, next\_state) is sampled from the replay memory.
  - **Target Calculation:** Target Q-values (`yi`) are calculated using the Bellman equation, considering immediate reward and the maximum Q-value for the next state.
  - **Loss Calculation:** The mean squared error loss is computed between predicted and target Q-values.

### 4. Hyperparameters

- **Batch Size ( `self.batch_size` ):** Size of mini-batches used during optimization.
- **Learning Rates ( `self.adamlr` , `self.targetlr` ):** Learning rates for the Adam optimizer used in the policy net.
- **Discount Factor ( `self.gamma` ):** Discount factor for future rewards in the Bellman equation.

In [8]:

```
class DQNAgent:
    def __init__(self):
        self.steps = 0
        self.memory = Memory(10000)

        self.batch_size = 256
        self.adamlr = 0.001
        self.gamma = 0.98

        # policy net, aka, Q function
        self.policy_net = BaseMLP(3, 9, self.adamlr)

    def action(self, state):

        with torch.no_grad():
            choice = float(torch.argmax(self.policy_net(state)).numpy())
            action = (choice - 4) / 2

        return choice, action

    def optimize(self):
        if len(self.memory) < self.batch_size:
            return

        batch = self.memory.sample(self.batch_size)
        state, action, reward, next_state = batch
        action = action.type(torch.int64)

        # target
        with torch.no_grad():    # parameters from previous iteration are held fixed
            yi = reward + self.gamma * self.policy_net(next_state).max(1)[0].unsqueeze(1)

        # Q(phi_i, a_i, theta)
        current_state_action = self.policy_net(state).gather(1, action)
```

```

# Back_propagation
loss = (yi - current_state_action) ** 2
self.policy_net.optimizer.zero_grad()
loss.mean().backward()
self.policy_net.optimizer.step()

```

## Training Setup

We train our model over a set number of episodes, where each episode represents a single game with a maximum of 200 steps. State updates occur after each step, and we check if the activity is completed.

To ensure an adequate amount of data for learning, we initiate backpropagation only after completing 1000 steps. This step threshold ensures a substantial dataset to sample from during the training process.

## Gym Environment Update

It's worth noting that in the updated Gym library, the environment now returns 5 values instead of 4. The 'done' signal has been divided into 'terminated' and 'truncated' to specify how the episode concluded. Currently, our model doesn't utilize this additional information.

## Logging

We take the cumulative sum of rewards as the key metric for this model. At the same time we also log the number of steps the model has sustained as an additional indicator of performance. After each training we also display a graph of the score change over the training period.

```

In [46]: def train_model(DQNAgentX, episodes):
    agent = DQNAgentX()
    env = gym.make('Pendulum-v1')
    score_list = []
    for episode in range(episodes):
        print(f"Episode: {episode}", end="")
        state, _ = env.reset()
        state = torch.tensor(state)
        done = False
        score = 0
        while not done:
            choice, action = agent.action(state)

            next_state, reward, terminated, truncated, _ = env.step([action])

            score += reward
            done = terminated or truncated

            if terminated:
                next_state = None
            else:
                next_state = torch.tensor(next_state, dtype=torch.float32)

            agent.memory.push((state, choice, reward, next_state))
            state = next_state

            if len(agent.memory) > 1000:
                agent.optimize()

        print(f", Score: {score}")
        score_list.append(score)

    env.close()

```

```
plt.plot(score_list)
plt.show()
```

```
return agent, score_list
```

```
In [20]: baseAgent, base_score = train_model(DQNAgent,400)
```

Episode: 0, Score: -1579.4195666572894  
Episode: 1, Score: -1603.513772210122  
Episode: 2, Score: -1485.4663338372372  
Episode: 3, Score: -1614.6608772658776  
Episode: 4, Score: -1665.7581725785196  
Episode: 5, Score: -1300.5187398415367  
Episode: 6, Score: -1439.800386187149  
Episode: 7, Score: -1268.57657133863  
Episode: 8, Score: -1503.0542223798968  
Episode: 9, Score: -1563.3456587971223  
Episode: 10, Score: -1236.465435257366  
Episode: 11, Score: -1188.8184146230865  
Episode: 12, Score: -1082.542962597973  
Episode: 13, Score: -1327.4170534703405  
Episode: 14, Score: -1012.4901028463876  
Episode: 15, Score: -1243.162094714658  
Episode: 16, Score: -970.7389728081843  
Episode: 17, Score: -1469.4509677322205  
Episode: 18, Score: -1000.1142904508289  
Episode: 19, Score: -1243.6948686697665  
Episode: 20, Score: -1242.3545577148557  
Episode: 21, Score: -1382.2397216814861  
Episode: 22, Score: -1217.305919466091  
Episode: 23, Score: -1244.6182254348903  
Episode: 24, Score: -1168.8360587798604  
Episode: 25, Score: -1271.4406996821242  
Episode: 26, Score: -904.2856811820027  
Episode: 27, Score: -1282.03569891004  
Episode: 28, Score: -976.9907649673164  
Episode: 29, Score: -978.7395613172323  
Episode: 30, Score: -1015.8981040452671  
Episode: 31, Score: -1046.2635625224023  
Episode: 32, Score: -1055.8020271335886  
Episode: 33, Score: -1157.074149501029  
Episode: 34, Score: -869.8697303608835  
Episode: 35, Score: -1088.4826199917877  
Episode: 36, Score: -1037.9806602242068  
Episode: 37, Score: -1022.7539167571592  
Episode: 38, Score: -1037.9211091277934  
Episode: 39, Score: -1023.0495689382597  
Episode: 40, Score: -976.2559789918341  
Episode: 41, Score: -857.5649605222625  
Episode: 42, Score: -1022.6765827961417  
Episode: 43, Score: -636.5697721548431  
Episode: 44, Score: -1126.0944671586542  
Episode: 45, Score: -873.7548717257127  
Episode: 46, Score: -835.5502632567703  
Episode: 47, Score: -866.9361786757045  
Episode: 48, Score: -752.1591199082923  
Episode: 49, Score: -872.8491521300881  
Episode: 50, Score: -862.3737128827413  
Episode: 51, Score: -1153.448998126797  
Episode: 52, Score: -877.0425812826049  
Episode: 53, Score: -976.0005081275016  
Episode: 54, Score: -505.0477043477925  
Episode: 55, Score: -501.6853349447782  
Episode: 56, Score: -130.26248025821954  
Episode: 57, Score: -862.0422034386407  
Episode: 58, Score: -374.6473964318185  
Episode: 59, Score: -830.3613776848587  
Episode: 60, Score: -619.3831488432434  
Episode: 61, Score: -495.57037512465666  
Episode: 62, Score: -720.2346276730929  
Episode: 63, Score: -368.1473580916557  
Episode: 64, Score: -699.3966342501021  
Episode: 65, Score: -612.3376563427912

Episode: 66, Score: -486.71302199861475  
Episode: 67, Score: -492.59301618665063  
Episode: 68, Score: -609.5879589727574  
Episode: 69, Score: -374.2074172927097  
Episode: 70, Score: -374.50809303666375  
Episode: 71, Score: -383.109038375067  
Episode: 72, Score: -360.5154568232899  
Episode: 73, Score: -394.8806371570027  
Episode: 74, Score: -128.4374051952028  
Episode: 75, Score: -117.96810923264006  
Episode: 76, Score: -356.8690950449388  
Episode: 77, Score: -1039.0937149776294  
Episode: 78, Score: -445.2823111607153  
Episode: 79, Score: -2.7292622605296284  
Episode: 80, Score: -1136.8001213229659  
Episode: 81, Score: -1117.66740015889  
Episode: 82, Score: -246.26644408569499  
Episode: 83, Score: -936.910266655586  
Episode: 84, Score: -1.287081485278547  
Episode: 85, Score: -127.04376834768564  
Episode: 86, Score: -831.6158075003483  
Episode: 87, Score: -1160.2195590117692  
Episode: 88, Score: -233.02224125662227  
Episode: 89, Score: -475.4619609757866  
Episode: 90, Score: -2.7075204759707416  
Episode: 91, Score: -120.36642376860844  
Episode: 92, Score: -127.07336084455783  
Episode: 93, Score: -124.00235269220414  
Episode: 94, Score: -283.22478176383214  
Episode: 95, Score: -244.06157550418345  
Episode: 96, Score: -259.14325956929855  
Episode: 97, Score: -119.74515551037504  
Episode: 98, Score: -1.5761797058935996  
Episode: 99, Score: -236.551058811127  
Episode: 100, Score: -130.62349108025822  
Episode: 101, Score: -129.97534236720622  
Episode: 102, Score: -2.4614255622538086  
Episode: 103, Score: -127.80720494479682  
Episode: 104, Score: -238.02108927064455  
Episode: 105, Score: -130.32144618817173  
Episode: 106, Score: -127.89281181597103  
Episode: 107, Score: -123.64004397577288  
Episode: 108, Score: -258.55269671429494  
Episode: 109, Score: -5.053454894099714  
Episode: 110, Score: -129.72757578377036  
Episode: 111, Score: -230.73318920191397  
Episode: 112, Score: -410.5637832889977  
Episode: 113, Score: -122.06215024102194  
Episode: 114, Score: -117.54895134315635  
Episode: 115, Score: -2.6606290542317863  
Episode: 116, Score: -117.94894030613321  
Episode: 117, Score: -123.77246685480354  
Episode: 118, Score: -127.78136577390623  
Episode: 119, Score: -324.2331584529673  
Episode: 120, Score: -248.40959356194998  
Episode: 121, Score: -122.08887743009811  
Episode: 122, Score: -130.06181736224002  
Episode: 123, Score: -379.75385545616  
Episode: 124, Score: -248.83895389478747  
Episode: 125, Score: -259.80097049723815  
Episode: 126, Score: -127.41644566532851  
Episode: 127, Score: -4.73655255007349  
Episode: 128, Score: -232.33284898585373  
Episode: 129, Score: -255.68415988059814  
Episode: 130, Score: -254.62638442984684  
Episode: 131, Score: -135.89782289768274

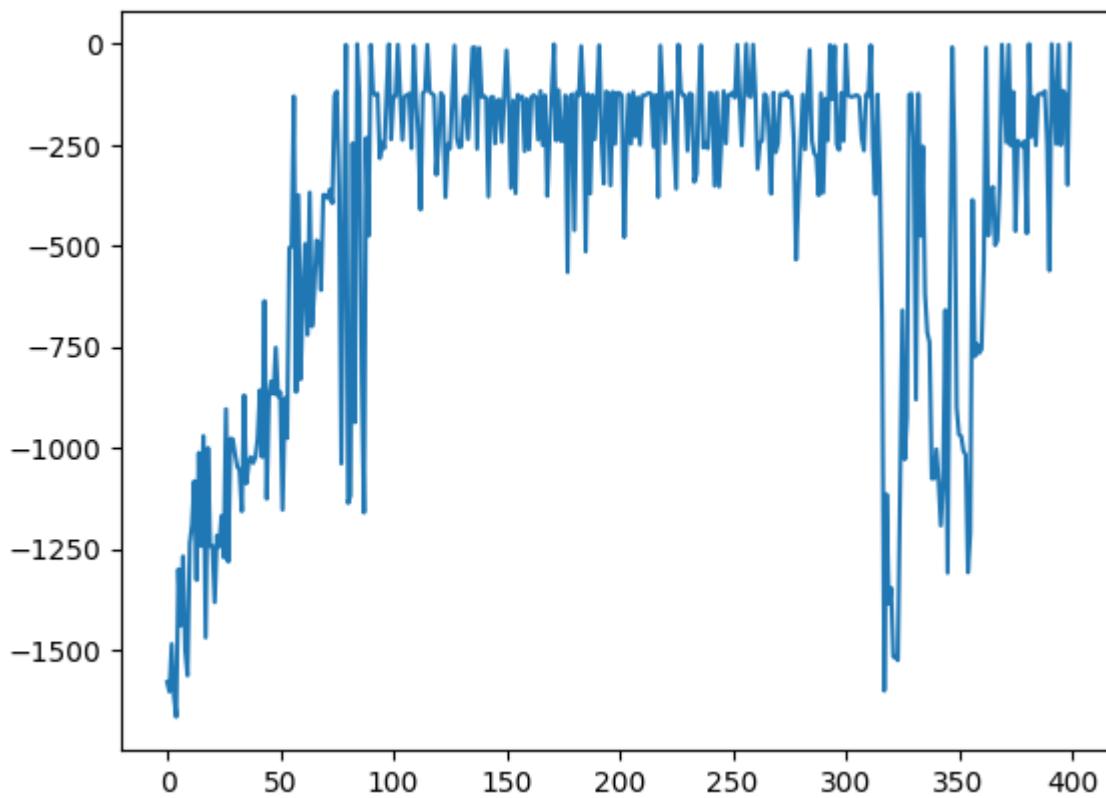
Episode: 132, Score: -129.1620459554708  
Episode: 133, Score: -236.11648421343546  
Episode: 134, Score: -131.642036876664  
Episode: 135, Score: -9.751415997166882  
Episode: 136, Score: -8.787252682605661  
Episode: 137, Score: -260.26715502824186  
Episode: 138, Score: -10.485453393038524  
Episode: 139, Score: -131.09146045086953  
Episode: 140, Score: -129.19706553896668  
Episode: 141, Score: -135.21391577436887  
Episode: 142, Score: -378.2230784469259  
Episode: 143, Score: -133.63661537350868  
Episode: 144, Score: -131.27301853356545  
Episode: 145, Score: -246.7032176945012  
Episode: 146, Score: -139.8360126295718  
Episode: 147, Score: -138.6548307187633  
Episode: 148, Score: -243.13764719587328  
Episode: 149, Score: -133.1222187086248  
Episode: 150, Score: -16.443920920091067  
Episode: 151, Score: -143.0151920531451  
Episode: 152, Score: -357.1563736014297  
Episode: 153, Score: -140.14221610329557  
Episode: 154, Score: -370.10037596747554  
Episode: 155, Score: -126.27599618826449  
Episode: 156, Score: -135.14528051657783  
Episode: 157, Score: -130.6554018278431  
Episode: 158, Score: -265.87742089971243  
Episode: 159, Score: -137.16049840687137  
Episode: 160, Score: -261.79258460150885  
Episode: 161, Score: -132.4704206080241  
Episode: 162, Score: -127.40167671212748  
Episode: 163, Score: -131.9040484063679  
Episode: 164, Score: -237.53923168665685  
Episode: 165, Score: -117.24898185138476  
Episode: 166, Score: -250.9876390154389  
Episode: 167, Score: -127.7106251157589  
Episode: 168, Score: -376.736219711012  
Episode: 169, Score: -234.57563047689118  
Episode: 170, Score: -124.02564226736179  
Episode: 171, Score: -1.521761315205473  
Episode: 172, Score: -239.5106596721319  
Episode: 173, Score: -114.80782635047568  
Episode: 174, Score: -240.76550801265154  
Episode: 175, Score: -240.24336616377948  
Episode: 176, Score: -127.9883940153761  
Episode: 177, Score: -565.4372749022596  
Episode: 178, Score: -131.54947455177648  
Episode: 179, Score: -125.3508197665462  
Episode: 180, Score: -461.58645885420873  
Episode: 181, Score: -120.48886341310934  
Episode: 182, Score: -128.41295174206616  
Episode: 183, Score: -6.01350601094747  
Episode: 184, Score: -124.8877285547786  
Episode: 185, Score: -514.3933252363902  
Episode: 186, Score: -124.33189188055496  
Episode: 187, Score: -370.875697680639  
Episode: 188, Score: -127.31402324312768  
Episode: 189, Score: -236.50335377674173  
Episode: 190, Score: -127.30224390718578  
Episode: 191, Score: -3.993585662059487  
Episode: 192, Score: -235.74495500455023  
Episode: 193, Score: -346.08342998982647  
Episode: 194, Score: -118.18896037202343  
Episode: 195, Score: -122.43832334764646  
Episode: 196, Score: -350.7337952057301  
Episode: 197, Score: -119.31459490205894

Episode: 198, Score: -248.08498625260322  
Episode: 199, Score: -121.45370603763455  
Episode: 200, Score: -122.34180283599095  
Episode: 201, Score: -128.43149837822793  
Episode: 202, Score: -479.2017577452237  
Episode: 203, Score: -259.42306981193826  
Episode: 204, Score: -125.49762363227649  
Episode: 205, Score: -247.45547119070648  
Episode: 206, Score: -119.47844829161208  
Episode: 207, Score: -229.7766219563504  
Episode: 208, Score: -132.78504276745656  
Episode: 209, Score: -249.4738155944694  
Episode: 210, Score: -128.41090052108922  
Episode: 211, Score: -128.222375404238  
Episode: 212, Score: -124.49814653616141  
Episode: 213, Score: -122.06206995364205  
Episode: 214, Score: -124.59480471309833  
Episode: 215, Score: -254.845056082065  
Episode: 216, Score: -126.91762764889349  
Episode: 217, Score: -379.505107617836  
Episode: 218, Score: -4.876749824632825  
Episode: 219, Score: -120.1791385629046  
Episode: 220, Score: -245.9750246159356  
Episode: 221, Score: -122.69137766646718  
Episode: 222, Score: -131.14713770386223  
Episode: 223, Score: -119.4689366021218  
Episode: 224, Score: -237.34728483695784  
Episode: 225, Score: -358.88934719060916  
Episode: 226, Score: -2.2634523500784356  
Episode: 227, Score: -115.03146533385467  
Episode: 228, Score: -128.39948950167192  
Episode: 229, Score: -128.7678198639756  
Episode: 230, Score: -264.5448983966691  
Episode: 231, Score: -123.07065588839419  
Episode: 232, Score: -125.99419983902128  
Episode: 233, Score: -342.21688614583957  
Episode: 234, Score: -322.6683129953942  
Episode: 235, Score: -125.59106227993725  
Episode: 236, Score: -4.24317729058369  
Episode: 237, Score: -257.4034268686337  
Episode: 238, Score: -245.0685789573449  
Episode: 239, Score: -259.1531574099767  
Episode: 240, Score: -121.62071280977274  
Episode: 241, Score: -124.86252349461722  
Episode: 242, Score: -351.6321652135742  
Episode: 243, Score: -120.71655298157296  
Episode: 244, Score: -353.5495741278746  
Episode: 245, Score: -232.94740049312546  
Episode: 246, Score: -117.3497894150093  
Episode: 247, Score: -246.8118084023975  
Episode: 248, Score: -126.49665753445156  
Episode: 249, Score: -126.78319402515237  
Episode: 250, Score: -120.72173656858385  
Episode: 251, Score: -133.83581677018265  
Episode: 252, Score: -2.637957046282448  
Episode: 253, Score: -127.74536458398151  
Episode: 254, Score: -251.68548726280616  
Episode: 255, Score: -132.9236657377012  
Episode: 256, Score: -1.0688642027306783  
Episode: 257, Score: -128.38090939798346  
Episode: 258, Score: -131.57060241105745  
Episode: 259, Score: -2.3789920884403597  
Episode: 260, Score: -122.82926515478339  
Episode: 261, Score: -309.52486842206474  
Episode: 262, Score: -244.93637334815486  
Episode: 263, Score: -239.94570629199373

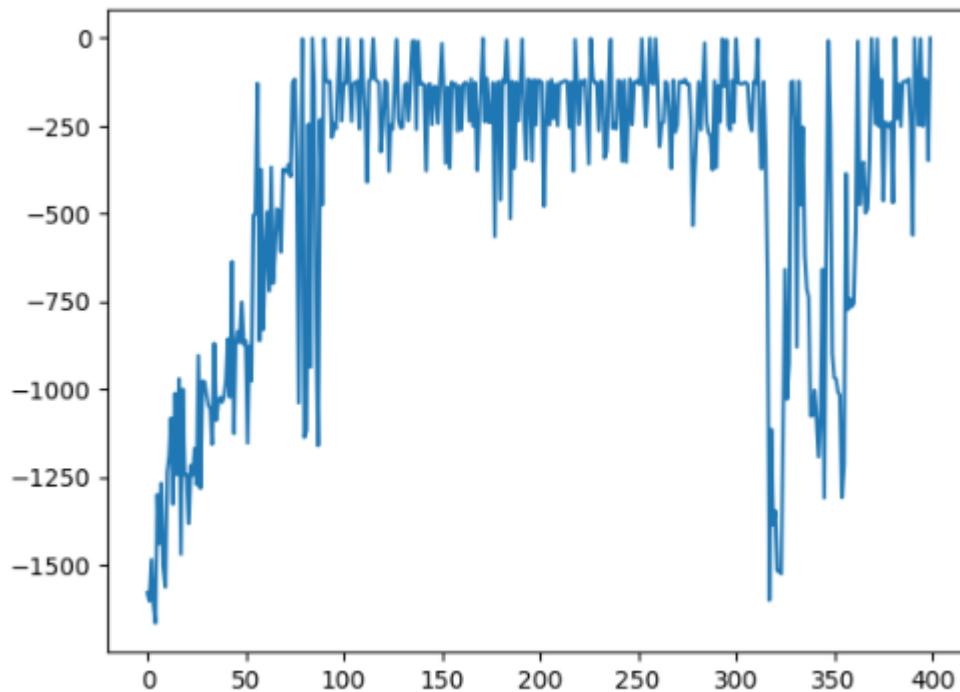
Episode: 264, Score: -124.76572947634551  
Episode: 265, Score: -131.74336401768136  
Episode: 266, Score: -231.88530826059562  
Episode: 267, Score: -371.30969001796495  
Episode: 268, Score: -121.2115774551211  
Episode: 269, Score: -268.5393382750411  
Episode: 270, Score: -245.96917465389348  
Episode: 271, Score: -124.97010228372652  
Episode: 272, Score: -124.86079099926839  
Episode: 273, Score: -127.52852086718254  
Episode: 274, Score: -118.87644717331878  
Episode: 275, Score: -132.10486004015758  
Episode: 276, Score: -131.42229986352336  
Episode: 277, Score: -236.75534608369694  
Episode: 278, Score: -534.2491609488625  
Episode: 279, Score: -376.86022399239846  
Episode: 280, Score: -255.00393288858746  
Episode: 281, Score: -125.57408429146612  
Episode: 282, Score: -261.4084698110387  
Episode: 283, Score: -135.03739980503445  
Episode: 284, Score: -15.054349750251967  
Episode: 285, Score: -239.99862155095062  
Episode: 286, Score: -271.00327745475045  
Episode: 287, Score: -279.371238775677  
Episode: 288, Score: -375.2749100727484  
Episode: 289, Score: -122.87107683483144  
Episode: 290, Score: -368.4853613365286  
Episode: 291, Score: -135.939243361402  
Episode: 292, Score: -239.09733482824902  
Episode: 293, Score: -3.382638842370028  
Episode: 294, Score: -137.35769528576756  
Episode: 295, Score: -5.726015515615015  
Episode: 296, Score: -249.31896750110081  
Episode: 297, Score: -262.84960185070827  
Episode: 298, Score: -121.9201597104326  
Episode: 299, Score: -240.4447344264919  
Episode: 300, Score: -3.058081592426182  
Episode: 301, Score: -127.33658900643064  
Episode: 302, Score: -130.8948627833655  
Episode: 303, Score: -131.23140402886605  
Episode: 304, Score: -128.513042004566  
Episode: 305, Score: -127.4559861045845  
Episode: 306, Score: -132.99538441503967  
Episode: 307, Score: -232.81528292919288  
Episode: 308, Score: -264.02202918555196  
Episode: 309, Score: -125.04195571928702  
Episode: 310, Score: -129.90120496000938  
Episode: 311, Score: -3.6653408879886085  
Episode: 312, Score: -257.68846891400483  
Episode: 313, Score: -372.53072829464327  
Episode: 314, Score: -125.78184645646539  
Episode: 315, Score: -314.4791495879945  
Episode: 316, Score: -681.7329912051232  
Episode: 317, Score: -1601.2034990082586  
Episode: 318, Score: -1115.0122250526524  
Episode: 319, Score: -1385.9393570365353  
Episode: 320, Score: -1346.5595532360906  
Episode: 321, Score: -1517.4037772773336  
Episode: 322, Score: -1516.1221300972786  
Episode: 323, Score: -1525.7951524482805  
Episode: 324, Score: -1095.6459781442832  
Episode: 325, Score: -659.8104031645689  
Episode: 326, Score: -1028.8154275955471  
Episode: 327, Score: -926.2473175439951  
Episode: 328, Score: -126.6493951711846  
Episode: 329, Score: -124.00556813458965

Episode: 330, Score: -236.96955353466578  
Episode: 331, Score: -880.5087305985973  
Episode: 332, Score: -123.85945000230657  
Episode: 333, Score: -475.621369339835  
Episode: 334, Score: -253.8417056193018  
Episode: 335, Score: -618.5448096915394  
Episode: 336, Score: -714.4264352983597  
Episode: 337, Score: -738.962364648034  
Episode: 338, Score: -1076.1130915789472  
Episode: 339, Score: -1075.860160481902  
Episode: 340, Score: -1003.5614719621652  
Episode: 341, Score: -1082.5239194075718  
Episode: 342, Score: -1192.3774944645377  
Episode: 343, Score: -1080.15244604252  
Episode: 344, Score: -659.1687893131975  
Episode: 345, Score: -1309.3769946540951  
Episode: 346, Score: -593.5629968522762  
Episode: 347, Score: -8.394422024728621  
Episode: 348, Score: -251.35687361219712  
Episode: 349, Score: -899.3578643728611  
Episode: 350, Score: -967.2320261450483  
Episode: 351, Score: -971.1599526656443  
Episode: 352, Score: -1007.6340200080027  
Episode: 353, Score: -1017.4348363158792  
Episode: 354, Score: -1308.149421986406  
Episode: 355, Score: -1211.2332450926276  
Episode: 356, Score: -386.6733498606246  
Episode: 357, Score: -773.8194633501934  
Episode: 358, Score: -743.1899943079412  
Episode: 359, Score: -765.3743255868885  
Episode: 360, Score: -755.7625777905073  
Episode: 361, Score: -545.8397011938098  
Episode: 362, Score: -9.797387999324727  
Episode: 363, Score: -474.61040518391115  
Episode: 364, Score: -367.0897801896093  
Episode: 365, Score: -353.93626773339423  
Episode: 366, Score: -498.80355938096386  
Episode: 367, Score: -486.5420253744084  
Episode: 368, Score: -346.92817101569926  
Episode: 369, Score: -2.186875580753638  
Episode: 370, Score: -120.2047804747041  
Episode: 371, Score: -244.3784061638022  
Episode: 372, Score: -2.985587569914941  
Episode: 373, Score: -252.377348575852  
Episode: 374, Score: -119.1336756609886  
Episode: 375, Score: -463.7777716816959  
Episode: 376, Score: -239.72592568412702  
Episode: 377, Score: -246.976687551602  
Episode: 378, Score: -253.89158083094415  
Episode: 379, Score: -240.72037510692758  
Episode: 380, Score: -469.4268721828124  
Episode: 381, Score: -0.7851404927241841  
Episode: 382, Score: -230.56537712585407  
Episode: 383, Score: -132.65084379959154  
Episode: 384, Score: -251.33194958330043  
Episode: 385, Score: -126.19608787316567  
Episode: 386, Score: -124.01609539082379  
Episode: 387, Score: -124.28682167595251  
Episode: 388, Score: -117.27734736817624  
Episode: 389, Score: -251.00034865211063  
Episode: 390, Score: -561.2801793365759  
Episode: 391, Score: -0.9262286198002831  
Episode: 392, Score: -125.9348755508311  
Episode: 393, Score: -248.4066311894941  
Episode: 394, Score: -2.7696993059906383  
Episode: 395, Score: -251.12637792581035

```
Episode: 396, Score: -116.21277884459721
Episode: 397, Score: -124.8328132785567
Episode: 398, Score: -349.43688935071543
Episode: 399, Score: -0.9955939609326678
```



```
In [65]: torch.save(baseAgent.policy_net.state_dict(), "./rl_models/base.pt")
showImage("./Img/RL/simple_dqn.png")
```



Observations:

- Model takes around 100 episodes to perform well
- Model is not very stable as it's score oscillates heavily throughout the training process even after 300 episodes
- Model is not consistent as you can see drastic decreases in score frequently. This is due to a inherent problem with vanilla DQNs called "Catastrophic Forgetting". This is because when a

model sees new unseen information and tries to optimise for it, it might forget previously learned information.

To define the performance of the model in a single metric, I will also compute the 100 epoch average.

```
In [12]: def evaluate_model(agent, episodes):

    # Initialize Pendulum environment for visualization
    env_visualize = gym.make('Pendulum-v1')
    scores = []
    # Visualization loop
    for episode in range(episodes):
        print(f"Evaluating Episode: {episode+1}", end="")

        # Reset the environment for a new episode
        state, info = env_visualize.reset()
        state = torch.tensor(state)

        done = False
        score = 0

        # Episode simulation loop
        while not done:
            # Render the environment (visualize the Pendulum)

            # Choose an action using the trained DQNAgent
            choice, action = agent.action(state)

            # Take a step in the environment
            next_state, reward, terminated, truncated, _ = env_visualize.step([action])

            # Update the total score
            score += reward
            done = terminated or truncated

            # If the episode is terminated, set next_state to None; otherwise, convert
            next_state = None if terminated else torch.tensor(next_state, dtype=torch.float)
            state = next_state

        scores.append(score)
        # Print episode score
        print(f", Score: {score}")
    print(f"Final Score: {np.mean(scores)}")
    # Close the visualization environment after visualizing
    env_visualize.close()
```

```
In [23]: evaluate_model(baseAgent, 100)
```

Evaluating Episode: 1, Score: -3.5438824587240347  
Evaluating Episode: 2, Score: -123.83347466187878  
Evaluating Episode: 3, Score: -121.02931821935753  
Evaluating Episode: 4, Score: -122.60711051593276  
Evaluating Episode: 5, Score: -123.87477125265588  
Evaluating Episode: 6, Score: -121.39073792407005  
Evaluating Episode: 7, Score: -232.45548308838042  
Evaluating Episode: 8, Score: -118.27427991300173  
Evaluating Episode: 9, Score: -1614.5584798730147  
Evaluating Episode: 10, Score: -125.9734429461688  
Evaluating Episode: 11, Score: -119.33248629284242  
Evaluating Episode: 12, Score: -225.7548796433445  
Evaluating Episode: 13, Score: -125.66170010634822  
Evaluating Episode: 14, Score: -232.49062314725506  
Evaluating Episode: 15, Score: -117.64351501030193  
Evaluating Episode: 16, Score: -119.42947160046955  
Evaluating Episode: 17, Score: -127.0705489585355  
Evaluating Episode: 18, Score: -123.747247490065  
Evaluating Episode: 19, Score: -122.49371005129927  
Evaluating Episode: 20, Score: -1.1974502301650078  
Evaluating Episode: 21, Score: -1611.3930311678887  
Evaluating Episode: 22, Score: -125.31308965094067  
Evaluating Episode: 23, Score: -122.53929855810125  
Evaluating Episode: 24, Score: -117.48910994803211  
Evaluating Episode: 25, Score: -125.33206680911579  
Evaluating Episode: 26, Score: -124.3829455610465  
Evaluating Episode: 27, Score: -0.5775502017606964  
Evaluating Episode: 28, Score: -124.97248001977302  
Evaluating Episode: 29, Score: -119.05037607630415  
Evaluating Episode: 30, Score: -1.1989125172779056  
Evaluating Episode: 31, Score: -126.6371779945372  
Evaluating Episode: 32, Score: -0.7745213056306275  
Evaluating Episode: 33, Score: -126.97031255337832  
Evaluating Episode: 34, Score: -361.67444401198424  
Evaluating Episode: 35, Score: -242.8750265267138  
Evaluating Episode: 36, Score: -116.60926173778357  
Evaluating Episode: 37, Score: -240.90136419628362  
Evaluating Episode: 38, Score: -262.9605893809454  
Evaluating Episode: 39, Score: -126.18607795400476  
Evaluating Episode: 40, Score: -128.59182730819347  
Evaluating Episode: 41, Score: -230.9883757495519  
Evaluating Episode: 42, Score: -352.1168248482551  
Evaluating Episode: 43, Score: -0.5762433705442921  
Evaluating Episode: 44, Score: -118.82985521490497  
Evaluating Episode: 45, Score: -125.75286950567705  
Evaluating Episode: 46, Score: -244.6623842080629  
Evaluating Episode: 47, Score: -261.483033153114  
Evaluating Episode: 48, Score: -117.04372989560656  
Evaluating Episode: 49, Score: -123.14058221361952  
Evaluating Episode: 50, Score: -1.4499977335256635  
Evaluating Episode: 51, Score: -374.45876026330615  
Evaluating Episode: 52, Score: -3.0342730144612315  
Evaluating Episode: 53, Score: -237.691626658912  
Evaluating Episode: 54, Score: -242.73911337794286  
Evaluating Episode: 55, Score: -247.9298262056673  
Evaluating Episode: 56, Score: -376.8257530889977  
Evaluating Episode: 57, Score: -114.1304780008841  
Evaluating Episode: 58, Score: -125.3077792755231  
Evaluating Episode: 59, Score: -120.40370966150661  
Evaluating Episode: 60, Score: -230.50296080454896  
Evaluating Episode: 61, Score: -125.99800659776605  
Evaluating Episode: 62, Score: -129.56913437902116  
Evaluating Episode: 63, Score: -122.92402400540195  
Evaluating Episode: 64, Score: -246.58460562149568  
Evaluating Episode: 65, Score: -122.24886488030262  
Evaluating Episode: 66, Score: -0.5448931711391756

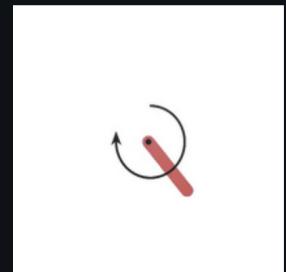
Evaluating Episode: 67, Score: -124.27015472680883  
Evaluating Episode: 68, Score: -371.0076240957725  
Evaluating Episode: 69, Score: -117.20684313654338  
Evaluating Episode: 70, Score: -235.80605911590925  
Evaluating Episode: 71, Score: -0.5074146809306377  
Evaluating Episode: 72, Score: -245.56625455451012  
Evaluating Episode: 73, Score: -130.42303420496438  
Evaluating Episode: 74, Score: -225.0051167862824  
Evaluating Episode: 75, Score: -2.4174011794134063  
Evaluating Episode: 76, Score: -125.59485452853869  
Evaluating Episode: 77, Score: -1.9691569180495039  
Evaluating Episode: 78, Score: -122.33828707357137  
Evaluating Episode: 79, Score: -126.04970360119314  
Evaluating Episode: 80, Score: -119.9905554571577  
Evaluating Episode: 81, Score: -1617.953582784215  
Evaluating Episode: 82, Score: -0.5583453915393884  
Evaluating Episode: 83, Score: -0.6692454739859495  
Evaluating Episode: 84, Score: -1622.111366644134  
Evaluating Episode: 85, Score: -127.0632714372061  
Evaluating Episode: 86, Score: -125.35814892359987  
Evaluating Episode: 87, Score: -226.874429163204  
Evaluating Episode: 88, Score: -1.0571426397926864  
Evaluating Episode: 89, Score: -248.93138647395145  
Evaluating Episode: 90, Score: -0.6003637728633788  
Evaluating Episode: 91, Score: -122.664394400944  
Evaluating Episode: 92, Score: -124.77279981923446  
Evaluating Episode: 93, Score: -1620.3909757619272  
Evaluating Episode: 94, Score: -126.58146291641322  
Evaluating Episode: 95, Score: -245.02708006692893  
Evaluating Episode: 96, Score: -126.32558832297109  
Evaluating Episode: 97, Score: -228.22916104586628  
Evaluating Episode: 98, Score: -116.36103220393856  
Evaluating Episode: 99, Score: -121.16191807893254  
Evaluating Episode: 100, Score: -127.09296454156873  
Final Score: -215.03666901686105

Score of -215 is not a bad score but I believe we can do better. I actually found a leaderboard from openai's repo. And the best score overall is -106. So our score of -215 is not bad at all.

## Pendulum-v0

The inverted pendulum swingup problem is a classic problem in the control literature. In this version of the problem, the pendulum starts in a random position, and the goal is to swing it up so it stays upright.

- [Environment details](#)
- *Pendulum-v0 is an unsolved environment, which means it does not have a specified reward threshold at which it's considered solved.*



| User                          | Best 100-episode performance | Write-up                          | Video                   |
|-------------------------------|------------------------------|-----------------------------------|-------------------------|
| <a href="#">KanishkNavale</a> | -106.9528                    | <a href="#">MultiAgent Policy</a> |                         |
| <a href="#">msinto93</a>      | -123.11 ± 6.86               | <a href="#">D4PG</a>              |                         |
| <a href="#">msinto93</a>      | -123.79 ± 6.90               | <a href="#">DDPG</a>              |                         |
| <a href="#">heerad</a>        | -134.48 ± 9.07               | <a href="#">writeup</a>           |                         |
| <a href="#">BS Haney</a>      | -135                         | <a href="#">Write-up</a>          | <a href="#">YouTube</a> |
| <a href="#">ThyrixYang</a>    | -136.16 ± 11.97              | <a href="#">writeup</a>           |                         |
| <a href="#">MaelFrancesc</a>  | -146.4 (mean 900 ep)         | <a href="#">writeup</a>           |                         |
| <a href="#">lirnli</a>        | -152.24 ± 10.87              | <a href="#">writeup</a>           |                         |

## Visualising Model Performance

While metrics are useful in viewing the performance of a model, sometimes seeing how it balances a pendulum can give insights into the performance of a model

```
In [25]: def visualize_model(agent, episodes):

    # Initialize Pendulum environment for visualization
    env_visualize = gym.make('Pendulum-v1', render_mode='human')

    # Visualization loop
    for episode in range(episodes):
        print(f"Visualizing Episode: {episode+1}", end="")

        # Reset the environment for a new episode
        state, info = env_visualize.reset()
        state = torch.tensor(state)

        done = False
        score = 0

        # Episode simulation loop
        while not done:
            # Render the environment (visualize the Pendulum)
            env_visualize.render()

            # Choose an action using the trained DQNAgent
            choice, action = agent.action(state)

            # Take a step in the environment
            next_state, reward, terminated, truncated, _ = env_visualize.step([action])

            # Update the total score
            score += reward
            done = terminated or truncated

            # If the episode is terminated, set next_state to None; otherwise, convert
            next_state = None if terminated else torch.tensor(next_state, dtype=torch
state = next_state

        for event in pygame.event.get():
            if event.type == pygame.QUIT:
                pygame.quit()
                sys.exit()

        # Print episode score
        print(f", Score: {score}")

    # Close the visualization environment after visualizing
    env_visualize.close()
```

```
In [26]: visualize_model(baseAgent,3)
```

```
Visualizing Episode: 1, Score: -221.7184656001412
Visualizing Episode: 2, Score: -117.59488499322866
Visualizing Episode: 3, Score: -126.89623739450441
```

Observations:

- Upon visual inspection, the model demonstrates the ability to moderately balance the pendulum.
- This suggests that the model has effectively learned to play the game well.

Issue:

- While the model successfully balances the pendulum on top, it is observed that it applies substantial force on both sides to maintain balance.
- The issue arises from the limitations of the action discrete array.
- The current array lacks fine-grained options near zero, preventing the model from exerting small amounts of force to delicately control the pendulum.

Solutions:

- Increase Action Array Size:
  - One potential solution is to increase the size of the action array, providing more options for the model to exert varying levels of force.
- Quadratic Spread:
  - Another approach is to make the spread of the action array quadratic, allowing for more nuanced control near zero.

By addressing the limitations of the action discrete array, the goal is to enable the model to exert precise control over the pendulum, leading to improved performance and a more realistic balance.

## Model Improvement

### Using Best Practices

Before moving on to more complicated improvements, I would like to get the basics right. I will implement this change to my model.

#### Leaky ReLU Activation:

- **Change:** Replace traditional ReLU activation functions with Leaky ReLU.
- **Reasoning:** Leaky ReLU allows a small, non-zero gradient for negative inputs, addressing the "dying ReLU" problem and providing a more robust learning signal during training.

This will involve changing the BaseMLP and DQNAgent Classes.

```
In [31]: class MLP(nn.Module):
    def __init__(self, inp, out, lr):
        super(MLP, self).__init__()

        self.layer1 = nn.Linear(inp, 32)
        self.layer2 = nn.Linear(32, 64)
        self.layer3 = nn.Linear(64, 32)
        self.layer4 = nn.Linear(32, out)

        self.lr = lr
        self.optimizer = optim.Adam(self.parameters(), lr=self.lr)

    def forward(self, x):
        x = self.layer1(x)
        x = F.leaky_relu(x)
        x = self.layer2(x)
        x = F.leaky_relu(x)
        x = self.layer3(x)
        x = F.leaky_relu(x)
        return self.layer4(x)
```

```
In [37]: class DQNAgentImp:
    def __init__(self):
        self.steps = 0
```

```

self.memory = Memory(10000)

self.batch_size = 256
self.adamlr = 0.001
self.gamma = 0.98

self.policy_net = MLP(3, 9, self.adamlr)

def action(self, state):

    with torch.no_grad():
        choice = float(torch.argmax(self.policy_net(state)).numpy())
        action = (choice - 4) / 2

    return choice, action

def optimize(self):
    if len(self.memory) < self.batch_size:
        return

    batch = self.memory.sample(self.batch_size)
    state, action, reward, next_state = batch
    action = action.type(torch.int64)

    # target
    with torch.no_grad():  # parameters from previous iteration are held fixed
        yi = reward + self.gamma * self.policy_net(next_state).max(1)[0].unsqueeze(-1)

    # Q(phi_i, a_i, theta)
    current_state_action = self.policy_net(state).gather(1, action)

    # Back_propagation
    loss = (yi - current_state_action) ** 2
    self.policy_net.optimizer.zero_grad()
    loss.mean().backward()
    self.policy_net.optimizer.step()

```

In [38]:

```

baseAgent, base_score = train_model(DQNAgent, 300)
improvedAgent, improved_score = train_model(DQNAgentImp, 300)
plot_tuples([(base_score, "base"), (improved_score, "best")], "Base vs Best Practices")

```

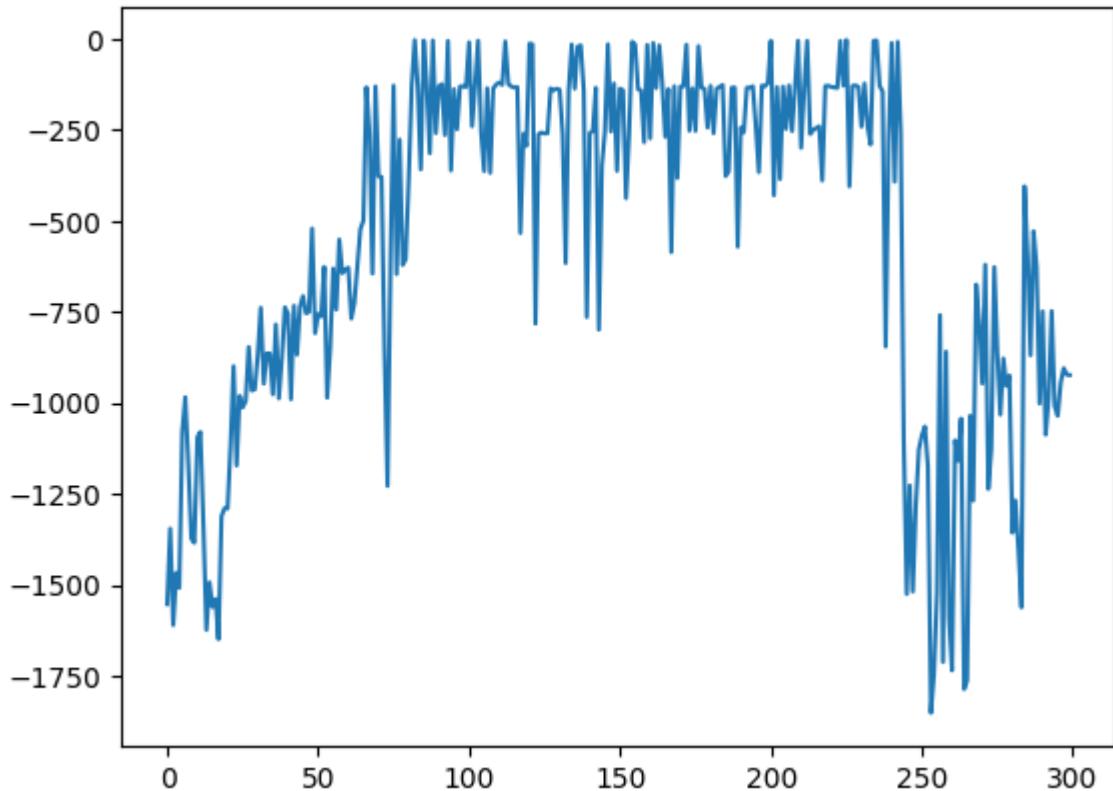
Episode: 0, Score: -1553.2304905499973  
Episode: 1, Score: -1345.786433827973  
Episode: 2, Score: -1610.8623846006146  
Episode: 3, Score: -1467.304045481317  
Episode: 4, Score: -1508.4749737380137  
Episode: 5, Score: -1079.4317592400578  
Episode: 6, Score: -984.5120400278432  
Episode: 7, Score: -1165.4430290780197  
Episode: 8, Score: -1372.9529315342947  
Episode: 9, Score: -1384.5929356151892  
Episode: 10, Score: -1094.9811055682258  
Episode: 11, Score: -1080.0286206677508  
Episode: 12, Score: -1336.1506245617513  
Episode: 13, Score: -1624.1738042134452  
Episode: 14, Score: -1493.1655109910896  
Episode: 15, Score: -1561.0438099877813  
Episode: 16, Score: -1539.0310030244348  
Episode: 17, Score: -1649.566996486605  
Episode: 18, Score: -1311.7490700303113  
Episode: 19, Score: -1289.4546604021882  
Episode: 20, Score: -1290.645230202009  
Episode: 21, Score: -1117.4564724205356  
Episode: 22, Score: -899.2375290251363  
Episode: 23, Score: -1172.2328140792215  
Episode: 24, Score: -980.8046883054704  
Episode: 25, Score: -1012.7012546302924  
Episode: 26, Score: -995.2309724167021  
Episode: 27, Score: -846.3600039584428  
Episode: 28, Score: -966.2293377483484  
Episode: 29, Score: -963.132129219019  
Episode: 30, Score: -869.8442177760018  
Episode: 31, Score: -738.7136821489945  
Episode: 32, Score: -947.243242206997  
Episode: 33, Score: -864.2247598047114  
Episode: 34, Score: -864.541988421007  
Episode: 35, Score: -976.6988318081335  
Episode: 36, Score: -784.6994156477843  
Episode: 37, Score: -987.7847537664051  
Episode: 38, Score: -875.6270606557247  
Episode: 39, Score: -737.4653444407444  
Episode: 40, Score: -755.2338934426276  
Episode: 41, Score: -990.1437925675903  
Episode: 42, Score: -733.1558213439006  
Episode: 43, Score: -867.7609803013247  
Episode: 44, Score: -735.5617994042282  
Episode: 45, Score: -706.4635189671221  
Episode: 46, Score: -754.0645090672211  
Episode: 47, Score: -749.8668243979521  
Episode: 48, Score: -520.8507918680912  
Episode: 49, Score: -808.4557437250464  
Episode: 50, Score: -755.8391823143596  
Episode: 51, Score: -762.7197197405898  
Episode: 52, Score: -626.4300895099707  
Episode: 53, Score: -985.7248115959779  
Episode: 54, Score: -850.1413116089059  
Episode: 55, Score: -630.888958926069  
Episode: 56, Score: -743.7172762747241  
Episode: 57, Score: -551.1470321423484  
Episode: 58, Score: -643.5904846449029  
Episode: 59, Score: -634.2954127059135  
Episode: 60, Score: -628.5188704590706  
Episode: 61, Score: -768.3648154161755  
Episode: 62, Score: -725.845108140222  
Episode: 63, Score: -623.6292346180484  
Episode: 64, Score: -522.5693909696475  
Episode: 65, Score: -499.08215305398943

Episode: 66, Score: -132.70369662304472  
Episode: 67, Score: -258.95441550459526  
Episode: 68, Score: -644.6651710631961  
Episode: 69, Score: -130.38861105416055  
Episode: 70, Score: -377.53989533149826  
Episode: 71, Score: -379.21322712283137  
Episode: 72, Score: -859.7118388497828  
Episode: 73, Score: -1228.5187660741212  
Episode: 74, Score: -606.8991002394325  
Episode: 75, Score: -128.3197938021031  
Episode: 76, Score: -646.1764426057423  
Episode: 77, Score: -276.5595791713747  
Episode: 78, Score: -621.0276189703143  
Episode: 79, Score: -605.1363392835458  
Episode: 80, Score: -387.01150248858136  
Episode: 81, Score: -128.80545491293609  
Episode: 82, Score: -3.0068963416815304  
Episode: 83, Score: -130.117152984996  
Episode: 84, Score: -358.3029099709864  
Episode: 85, Score: -4.430587403856548  
Episode: 86, Score: -125.8726582675238  
Episode: 87, Score: -314.02453202025134  
Episode: 88, Score: -3.995033350104028  
Episode: 89, Score: -258.62648492752726  
Episode: 90, Score: -129.62765317839262  
Episode: 91, Score: -124.23857537782649  
Episode: 92, Score: -263.8725505867858  
Episode: 93, Score: -5.2928234171700534  
Episode: 94, Score: -361.2203962012073  
Episode: 95, Score: -137.8281969710115  
Episode: 96, Score: -248.21779382997246  
Episode: 97, Score: -130.42847574849029  
Episode: 98, Score: -130.09185685960048  
Episode: 99, Score: -131.85928356406978  
Episode: 100, Score: -9.081281875121642  
Episode: 101, Score: -239.7252499786899  
Episode: 102, Score: -126.50444435338066  
Episode: 103, Score: -4.690328254934017  
Episode: 104, Score: -256.39783851557536  
Episode: 105, Score: -363.2112717477336  
Episode: 106, Score: -135.14612599978133  
Episode: 107, Score: -367.85494588065455  
Episode: 108, Score: -136.26990874317352  
Episode: 109, Score: -125.00435141188379  
Episode: 110, Score: -119.8681757446986  
Episode: 111, Score: -125.91451724981951  
Episode: 112, Score: -6.770567811979129  
Episode: 113, Score: -123.75394515721734  
Episode: 114, Score: -130.3124235713384  
Episode: 115, Score: -133.48949808310135  
Episode: 116, Score: -132.1423326808043  
Episode: 117, Score: -533.3259096769049  
Episode: 118, Score: -260.7407677524274  
Episode: 119, Score: -293.86627752676  
Episode: 120, Score: -12.632858660903402  
Episode: 121, Score: -14.608216231540815  
Episode: 122, Score: -782.4708073790741  
Episode: 123, Score: -261.70996314954783  
Episode: 124, Score: -257.2419666499545  
Episode: 125, Score: -259.82972930680205  
Episode: 126, Score: -259.0410030539339  
Episode: 127, Score: -135.83236256617803  
Episode: 128, Score: -140.9605675771004  
Episode: 129, Score: -136.59728482923055  
Episode: 130, Score: -139.49056257441117  
Episode: 131, Score: -260.82424055676

Episode: 132, Score: -615.8240627931406  
Episode: 133, Score: -141.1754934107095  
Episode: 134, Score: -14.955371689196385  
Episode: 135, Score: -136.39338733547027  
Episode: 136, Score: -20.919520205396797  
Episode: 137, Score: -17.704742337838542  
Episode: 138, Score: -141.5608079118641  
Episode: 139, Score: -764.2122617321595  
Episode: 140, Score: -258.46476409501497  
Episode: 141, Score: -253.48209947139873  
Episode: 142, Score: -134.37783228023375  
Episode: 143, Score: -798.9179917391833  
Episode: 144, Score: -348.87072916493645  
Episode: 145, Score: -254.18858369287483  
Episode: 146, Score: -14.368644039997  
Episode: 147, Score: -254.4962595179025  
Episode: 148, Score: -122.41743133162618  
Episode: 149, Score: -362.46073853789625  
Episode: 150, Score: -136.5438578120226  
Episode: 151, Score: -141.14482102128036  
Episode: 152, Score: -437.60751744460157  
Episode: 153, Score: -263.6865239267191  
Episode: 154, Score: -7.723493904007846  
Episode: 155, Score: -15.3341953449061  
Episode: 156, Score: -138.17065400319916  
Episode: 157, Score: -141.27878115666553  
Episode: 158, Score: -283.72676054583405  
Episode: 159, Score: -16.341334114250767  
Episode: 160, Score: -273.23525028215636  
Episode: 161, Score: -11.147778371450084  
Episode: 162, Score: -134.8591676501316  
Episode: 163, Score: -17.55447661883781  
Episode: 164, Score: -121.75835828103371  
Episode: 165, Score: -268.63455592589975  
Episode: 166, Score: -137.5765775332735  
Episode: 167, Score: -586.0073504641947  
Episode: 168, Score: -129.832112724978  
Episode: 169, Score: -382.23700914113795  
Episode: 170, Score: -130.29151934937389  
Episode: 171, Score: -130.73915899961023  
Episode: 172, Score: -15.735730526354196  
Episode: 173, Score: -252.1364461777315  
Episode: 174, Score: -135.75836441295672  
Episode: 175, Score: -252.62258097175143  
Episode: 176, Score: -19.510048524248305  
Episode: 177, Score: -133.78319485415773  
Episode: 178, Score: -136.90357708925308  
Episode: 179, Score: -243.5442163496364  
Episode: 180, Score: -128.75401192225743  
Episode: 181, Score: -258.8072849779167  
Episode: 182, Score: -136.78735267277003  
Episode: 183, Score: -131.4417245068468  
Episode: 184, Score: -126.38638753707093  
Episode: 185, Score: -376.7976523931247  
Episode: 186, Score: -363.89850730255904  
Episode: 187, Score: -133.98067591310522  
Episode: 188, Score: -133.39791630526275  
Episode: 189, Score: -570.9587581477624  
Episode: 190, Score: -243.41748406885804  
Episode: 191, Score: -256.62736579510454  
Episode: 192, Score: -134.10246736188432  
Episode: 193, Score: -133.8854858468728  
Episode: 194, Score: -130.08625030285154  
Episode: 195, Score: -239.68665047441456  
Episode: 196, Score: -365.98315677794983  
Episode: 197, Score: -128.0411192793648

Episode: 198, Score: -130.07084938448256  
Episode: 199, Score: -120.39497174209191  
Episode: 200, Score: -4.891312660460766  
Episode: 201, Score: -429.522561547655  
Episode: 202, Score: -132.35528185379653  
Episode: 203, Score: -385.08869058911193  
Episode: 204, Score: -130.81644651125436  
Episode: 205, Score: -247.31347265970567  
Episode: 206, Score: -127.88013879237221  
Episode: 207, Score: -253.1638472334902  
Episode: 208, Score: -129.12240576406438  
Episode: 209, Score: -4.786701484443907  
Episode: 210, Score: -299.1300093880408  
Episode: 211, Score: -133.40321346722303  
Episode: 212, Score: -5.025958902259312  
Episode: 213, Score: -261.25204139239816  
Episode: 214, Score: -249.9367671782165  
Episode: 215, Score: -245.0622561099887  
Episode: 216, Score: -240.0067948531571  
Episode: 217, Score: -388.8708947668383  
Episode: 218, Score: -129.41324556815024  
Episode: 219, Score: -128.75992496080426  
Episode: 220, Score: -131.60766501018955  
Episode: 221, Score: -133.42919597193708  
Episode: 222, Score: -134.48211300987313  
Episode: 223, Score: -5.277135589390699  
Episode: 224, Score: -128.21432177956595  
Episode: 225, Score: -2.841696043937849  
Episode: 226, Score: -405.07265118014067  
Episode: 227, Score: -129.0153445640581  
Episode: 228, Score: -127.61515722988253  
Episode: 229, Score: -131.8821135812209  
Episode: 230, Score: -241.0405154833229  
Episode: 231, Score: -122.363924576342  
Episode: 232, Score: -238.96710471548374  
Episode: 233, Score: -289.8893444854551  
Episode: 234, Score: -5.276431712367982  
Episode: 235, Score: -4.104726025448944  
Episode: 236, Score: -128.40055020608372  
Episode: 237, Score: -144.61772482058893  
Episode: 238, Score: -845.3561013751594  
Episode: 239, Score: -277.65099539948505  
Episode: 240, Score: -11.000118956149471  
Episode: 241, Score: -392.305812542984  
Episode: 242, Score: -7.8569424553551235  
Episode: 243, Score: -256.24776996289756  
Episode: 244, Score: -1103.5266298658125  
Episode: 245, Score: -1525.6868324472748  
Episode: 246, Score: -1227.1876239762034  
Episode: 247, Score: -1518.7194304738005  
Episode: 248, Score: -1281.6900484770335  
Episode: 249, Score: -1128.264702287209  
Episode: 250, Score: -1089.9644015561985  
Episode: 251, Score: -1065.1820541494262  
Episode: 252, Score: -1174.2389677939013  
Episode: 253, Score: -1852.0379799421753  
Episode: 254, Score: -1745.3550400307015  
Episode: 255, Score: -1521.495020221132  
Episode: 256, Score: -758.9262024415282  
Episode: 257, Score: -1712.3342662262173  
Episode: 258, Score: -859.239490434705  
Episode: 259, Score: -1590.4532360212525  
Episode: 260, Score: -1735.2927363238368  
Episode: 261, Score: -1103.4072670580713  
Episode: 262, Score: -1158.9979013882455  
Episode: 263, Score: -1043.2450389678527

Episode: 264, Score: -1787.0646149825554  
Episode: 265, Score: -1762.8393471254374  
Episode: 266, Score: -1035.3149679019252  
Episode: 267, Score: -1268.929245490723  
Episode: 268, Score: -675.0509371067047  
Episode: 269, Score: -780.6927314504618  
Episode: 270, Score: -946.7917526418597  
Episode: 271, Score: -620.3426553766444  
Episode: 272, Score: -1236.6133485879982  
Episode: 273, Score: -1126.6811501899324  
Episode: 274, Score: -626.9328081694163  
Episode: 275, Score: -861.5529610965231  
Episode: 276, Score: -1032.4524354999521  
Episode: 277, Score: -878.071569957542  
Episode: 278, Score: -954.0387064745223  
Episode: 279, Score: -924.9414174576151  
Episode: 280, Score: -1356.7797308030488  
Episode: 281, Score: -1268.2540422767568  
Episode: 282, Score: -1405.076109861573  
Episode: 283, Score: -1562.2074063229002  
Episode: 284, Score: -405.07555330611626  
Episode: 285, Score: -593.8231608618735  
Episode: 286, Score: -869.230730692613  
Episode: 287, Score: -528.4106059020986  
Episode: 288, Score: -626.0889333877358  
Episode: 289, Score: -1002.0711513568383  
Episode: 290, Score: -748.6904062592226  
Episode: 291, Score: -1087.4368130677399  
Episode: 292, Score: -986.7777170740212  
Episode: 293, Score: -747.7837286423969  
Episode: 294, Score: -1009.2093106956748  
Episode: 295, Score: -1035.7722432406185  
Episode: 296, Score: -944.150898687658  
Episode: 297, Score: -904.7484737806642  
Episode: 298, Score: -922.7073643518416  
Episode: 299, Score: -923.652839445085



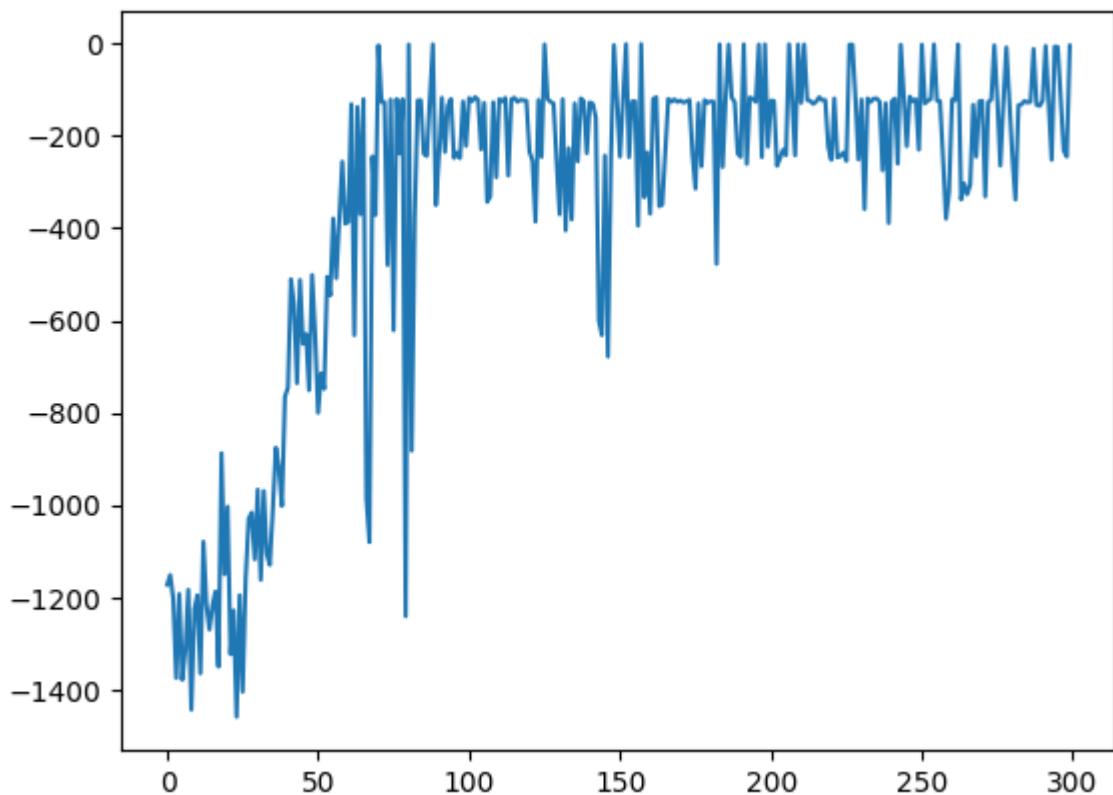
Episode: 0, Score: -1172.9617595045681  
Episode: 1, Score: -1151.9078940643892  
Episode: 2, Score: -1202.9726589044662  
Episode: 3, Score: -1375.338247208075  
Episode: 4, Score: -1193.1050051442926  
Episode: 5, Score: -1379.361256233199  
Episode: 6, Score: -1314.922571886764  
Episode: 7, Score: -1183.9721651189222  
Episode: 8, Score: -1444.0077554311779  
Episode: 9, Score: -1226.3325619801828  
Episode: 10, Score: -1196.1173116245566  
Episode: 11, Score: -1365.2837073861992  
Episode: 12, Score: -1079.7076248015692  
Episode: 13, Score: -1215.7502069795764  
Episode: 14, Score: -1270.234252758934  
Episode: 15, Score: -1217.7533896741063  
Episode: 16, Score: -1187.7903343172832  
Episode: 17, Score: -1350.6759232861016  
Episode: 18, Score: -887.7760896008308  
Episode: 19, Score: -1150.3970110838623  
Episode: 20, Score: -1003.7077000242513  
Episode: 21, Score: -1323.6849422253572  
Episode: 22, Score: -1228.784805143345  
Episode: 23, Score: -1459.4790462149845  
Episode: 24, Score: -1195.7312187902487  
Episode: 25, Score: -1405.1736730947607  
Episode: 26, Score: -1158.7063266293023  
Episode: 27, Score: -1029.8364183406454  
Episode: 28, Score: -1016.7066193394087  
Episode: 29, Score: -1117.941586346379  
Episode: 30, Score: -966.7758200350683  
Episode: 31, Score: -1162.4902925941549  
Episode: 32, Score: -969.7971414273904  
Episode: 33, Score: -1106.5286577233746  
Episode: 34, Score: -1129.9048323259979  
Episode: 35, Score: -1008.5068417807832  
Episode: 36, Score: -875.8420303242574  
Episode: 37, Score: -919.5224696004113  
Episode: 38, Score: -1002.3251079179493  
Episode: 39, Score: -765.8086202075306  
Episode: 40, Score: -745.6187541641999  
Episode: 41, Score: -510.6687865031044  
Episode: 42, Score: -565.0714646327696  
Episode: 43, Score: -736.0967429395346  
Episode: 44, Score: -512.0853037229488  
Episode: 45, Score: -650.7139871490889  
Episode: 46, Score: -629.5100150689088  
Episode: 47, Score: -751.3740215263408  
Episode: 48, Score: -501.2359117264534  
Episode: 49, Score: -628.9246917167661  
Episode: 50, Score: -800.1938425819517  
Episode: 51, Score: -714.8191433765801  
Episode: 52, Score: -748.281270545106  
Episode: 53, Score: -505.59703372586307  
Episode: 54, Score: -545.9184406687671  
Episode: 55, Score: -378.8660829290605  
Episode: 56, Score: -508.49027623992663  
Episode: 57, Score: -380.6125442201713  
Episode: 58, Score: -255.7302816910776  
Episode: 59, Score: -390.3681624282863  
Episode: 60, Score: -385.9873715747481  
Episode: 61, Score: -130.89395391394805  
Episode: 62, Score: -631.960689901474  
Episode: 63, Score: -137.0981365772906  
Episode: 64, Score: -369.85219249793477  
Episode: 65, Score: -120.24678291058824

Episode: 66, Score: -986.4697086043379  
Episode: 67, Score: -1080.369529460529  
Episode: 68, Score: -244.39892961973715  
Episode: 69, Score: -370.998243770006  
Episode: 70, Score: -3.488769068429411  
Episode: 71, Score: -126.55870087802343  
Episode: 72, Score: -126.3658914698142  
Episode: 73, Score: -479.84083407207146  
Episode: 74, Score: -121.1730221362277  
Episode: 75, Score: -621.4945356152815  
Episode: 76, Score: -119.77191142050934  
Episode: 77, Score: -238.3230815547707  
Episode: 78, Score: -120.04252026884359  
Episode: 79, Score: -1241.336648192486  
Episode: 80, Score: -1.7042328563064388  
Episode: 81, Score: -882.5053799135226  
Episode: 82, Score: -372.91800042353424  
Episode: 83, Score: -122.99069844943024  
Episode: 84, Score: -122.48492038237409  
Episode: 85, Score: -238.50878168878643  
Episode: 86, Score: -243.41680372778322  
Episode: 87, Score: -124.81700743667406  
Episode: 88, Score: -1.2792808095722368  
Episode: 89, Score: -349.89739605548937  
Episode: 90, Score: -244.95523660332327  
Episode: 91, Score: -117.24431796137874  
Episode: 92, Score: -234.78614896396542  
Episode: 93, Score: -134.64715412399042  
Episode: 94, Score: -119.59914940592377  
Episode: 95, Score: -247.89869475712487  
Episode: 96, Score: -235.7524853695198  
Episode: 97, Score: -248.23148710386988  
Episode: 98, Score: -125.2193626264223  
Episode: 99, Score: -221.63584020004538  
Episode: 100, Score: -117.7049842749142  
Episode: 101, Score: -124.69089609660728  
Episode: 102, Score: -115.74750465785496  
Episode: 103, Score: -121.75642851306787  
Episode: 104, Score: -229.1434110449737  
Episode: 105, Score: -128.32075405509926  
Episode: 106, Score: -343.01834449057816  
Episode: 107, Score: -331.637900006163  
Episode: 108, Score: -126.8530813317099  
Episode: 109, Score: -289.4855965289848  
Episode: 110, Score: -120.21054801640524  
Episode: 111, Score: -125.54371664756154  
Episode: 112, Score: -117.45350312892056  
Episode: 113, Score: -285.7752263739603  
Episode: 114, Score: -120.92089514276411  
Episode: 115, Score: -117.13798365071007  
Episode: 116, Score: -126.2072592093211  
Episode: 117, Score: -121.60399161068769  
Episode: 118, Score: -122.53153982724515  
Episode: 119, Score: -123.53771074054255  
Episode: 120, Score: -233.4257011417581  
Episode: 121, Score: -253.69001810503374  
Episode: 122, Score: -385.79311151626956  
Episode: 123, Score: -121.94028145962835  
Episode: 124, Score: -245.5115654235085  
Episode: 125, Score: -1.4111064677211667  
Episode: 126, Score: -120.0113938999202  
Episode: 127, Score: -125.89843543931707  
Episode: 128, Score: -129.84729664773513  
Episode: 129, Score: -255.4331254928001  
Episode: 130, Score: -369.7578920021486  
Episode: 131, Score: -120.31598962040475

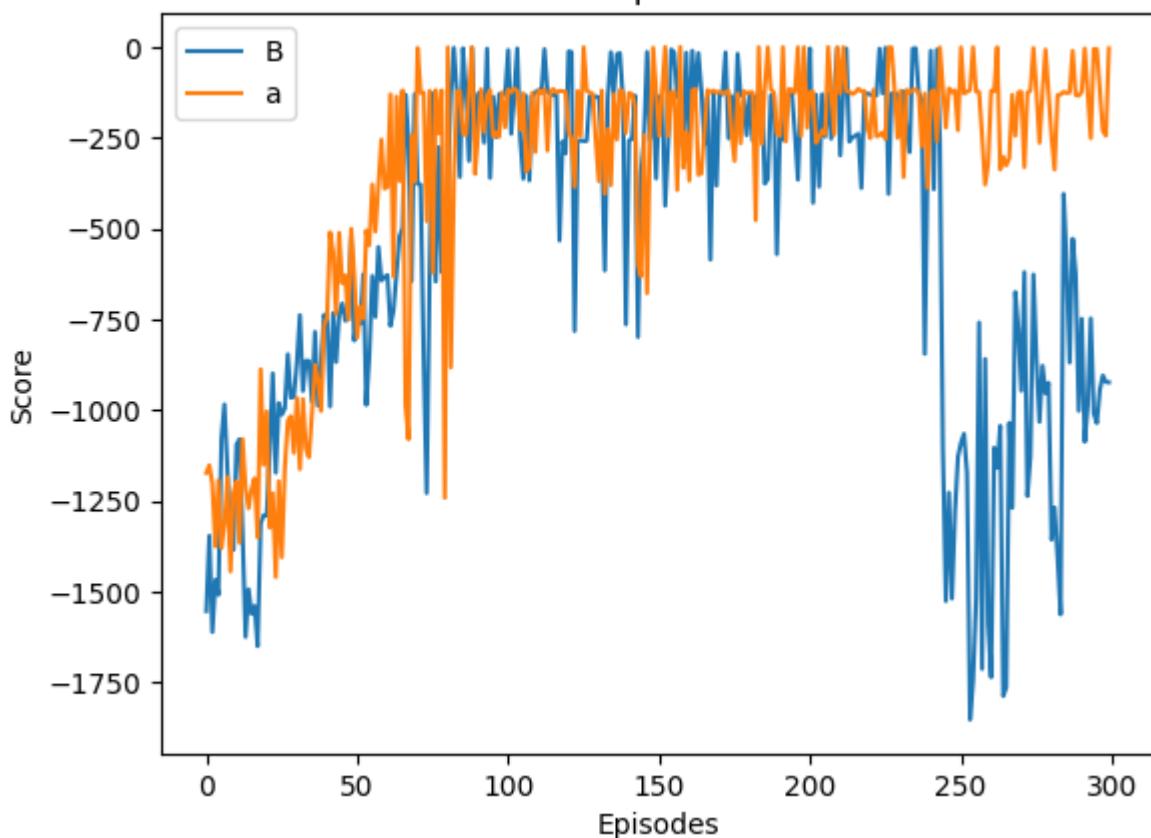
Episode: 132, Score: -404.75226178953  
Episode: 133, Score: -226.5866353013665  
Episode: 134, Score: -381.0070586660787  
Episode: 135, Score: -129.51566617591044  
Episode: 136, Score: -255.1670847090436  
Episode: 137, Score: -118.68557819957859  
Episode: 138, Score: -124.46692785533162  
Episode: 139, Score: -237.25890850804825  
Episode: 140, Score: -127.33928174195337  
Episode: 141, Score: -130.35942480465468  
Episode: 142, Score: -160.5562292781752  
Episode: 143, Score: -600.9331898281511  
Episode: 144, Score: -631.9444517623706  
Episode: 145, Score: -242.7379282706336  
Episode: 146, Score: -678.2794415507663  
Episode: 147, Score: -247.60820259202293  
Episode: 148, Score: -2.917393960028613  
Episode: 149, Score: -125.37685559341661  
Episode: 150, Score: -244.8590907893427  
Episode: 151, Score: -119.19265201451152  
Episode: 152, Score: -0.7839806550762325  
Episode: 153, Score: -245.92402502847872  
Episode: 154, Score: -125.11776735816176  
Episode: 155, Score: -124.65036175952459  
Episode: 156, Score: -394.4547546053306  
Episode: 157, Score: -0.6401406696180749  
Episode: 158, Score: -333.05516484847766  
Episode: 159, Score: -235.44997405048306  
Episode: 160, Score: -368.5804514564257  
Episode: 161, Score: -119.96480650771451  
Episode: 162, Score: -115.96123139436966  
Episode: 163, Score: -352.1258636963924  
Episode: 164, Score: -349.03007411352206  
Episode: 165, Score: -237.64914833646455  
Episode: 166, Score: -119.55963409302792  
Episode: 167, Score: -125.95793638434712  
Episode: 168, Score: -120.61404389455161  
Episode: 169, Score: -126.14944013227782  
Episode: 170, Score: -123.36980776871101  
Episode: 171, Score: -127.5052099645253  
Episode: 172, Score: -125.52571695978418  
Episode: 173, Score: -121.6606581998289  
Episode: 174, Score: -237.61227308412714  
Episode: 175, Score: -314.2376065848769  
Episode: 176, Score: -129.33964128683996  
Episode: 177, Score: -265.7364462505695  
Episode: 178, Score: -122.26569245184389  
Episode: 179, Score: -127.85286079596729  
Episode: 180, Score: -125.32538637882857  
Episode: 181, Score: -125.62114781534832  
Episode: 182, Score: -477.3990882746743  
Episode: 183, Score: -1.267178087135147  
Episode: 184, Score: -267.1009628765066  
Episode: 185, Score: -117.25673765085003  
Episode: 186, Score: -1.8141369394073308  
Episode: 187, Score: -116.27982619045403  
Episode: 188, Score: -126.6601368574768  
Episode: 189, Score: -238.29497087302232  
Episode: 190, Score: -245.37832120267348  
Episode: 191, Score: -1.216235291161834  
Episode: 192, Score: -260.64067030974275  
Episode: 193, Score: -116.44724158495922  
Episode: 194, Score: -119.42910934133882  
Episode: 195, Score: -126.2178237490941  
Episode: 196, Score: -1.6934843654263159  
Episode: 197, Score: -246.04740111621288

Episode: 198, Score: -1.3382744522963277  
Episode: 199, Score: -223.22332332246526  
Episode: 200, Score: -123.32723104913241  
Episode: 201, Score: -124.11730180591604  
Episode: 202, Score: -264.7765788872812  
Episode: 203, Score: -248.6062710814045  
Episode: 204, Score: -229.9998261667538  
Episode: 205, Score: -241.83514541563633  
Episode: 206, Score: -0.9941322970792021  
Episode: 207, Score: -121.22652072646274  
Episode: 208, Score: -242.0177504196922  
Episode: 209, Score: -1.5286871505460813  
Episode: 210, Score: -128.22052424139736  
Episode: 211, Score: -1.559880918041513  
Episode: 212, Score: -121.85584044090618  
Episode: 213, Score: -123.73514491416978  
Episode: 214, Score: -130.43485405452878  
Episode: 215, Score: -125.42875324185985  
Episode: 216, Score: -115.86949746736705  
Episode: 217, Score: -122.49857524482209  
Episode: 218, Score: -122.0965108046342  
Episode: 219, Score: -223.50888152013545  
Episode: 220, Score: -251.7500070860728  
Episode: 221, Score: -118.61893499309879  
Episode: 222, Score: -246.60989783882707  
Episode: 223, Score: -243.66861527496604  
Episode: 224, Score: -235.94933011353186  
Episode: 225, Score: -254.10573853615065  
Episode: 226, Score: -1.9920148632552355  
Episode: 227, Score: -1.9975802840155508  
Episode: 228, Score: -119.04469734694321  
Episode: 229, Score: -251.03998495400057  
Episode: 230, Score: -121.01012972164244  
Episode: 231, Score: -358.73872446288607  
Episode: 232, Score: -118.67827832186413  
Episode: 233, Score: -126.48929851554718  
Episode: 234, Score: -119.18277043358978  
Episode: 235, Score: -118.50693113032044  
Episode: 236, Score: -125.59744128363891  
Episode: 237, Score: -274.55849639661307  
Episode: 238, Score: -129.0941666637343  
Episode: 239, Score: -388.7603519050029  
Episode: 240, Score: -125.61908017895132  
Episode: 241, Score: -119.0620136405465  
Episode: 242, Score: -260.0184031005041  
Episode: 243, Score: -2.1759079300698803  
Episode: 244, Score: -115.58035499621762  
Episode: 245, Score: -222.29916931209237  
Episode: 246, Score: -114.66386056505837  
Episode: 247, Score: -125.11712195747477  
Episode: 248, Score: -120.49608052135743  
Episode: 249, Score: -229.7645817331241  
Episode: 250, Score: -1.9460617542404997  
Episode: 251, Score: -129.32766282018005  
Episode: 252, Score: -123.45264206589708  
Episode: 253, Score: -119.58081835910124  
Episode: 254, Score: -1.452571770538841  
Episode: 255, Score: -123.32021600335086  
Episode: 256, Score: -124.4870752696748  
Episode: 257, Score: -245.86995473459592  
Episode: 258, Score: -379.2695270682616  
Episode: 259, Score: -321.14714393851085  
Episode: 260, Score: -121.14971874933549  
Episode: 261, Score: -122.69143703337731  
Episode: 262, Score: -1.5468055443814153  
Episode: 263, Score: -337.7322827061102

Episode: 264, Score: -302.9584806667882  
Episode: 265, Score: -326.6407872055476  
Episode: 266, Score: -306.0565617266078  
Episode: 267, Score: -132.24295552702824  
Episode: 268, Score: -244.85010465609156  
Episode: 269, Score: -125.20096213480358  
Episode: 270, Score: -123.79976171795296  
Episode: 271, Score: -331.2953516868401  
Episode: 272, Score: -126.5405328965086  
Episode: 273, Score: -120.94795584318858  
Episode: 274, Score: -3.4867033810156105  
Episode: 275, Score: -121.78880924046311  
Episode: 276, Score: -264.38711534614754  
Episode: 277, Score: -127.52565674454583  
Episode: 278, Score: -8.011604923501196  
Episode: 279, Score: -129.96745776601426  
Episode: 280, Score: -245.88710111078595  
Episode: 281, Score: -337.9857070721282  
Episode: 282, Score: -133.32003677147304  
Episode: 283, Score: -131.48521000590338  
Episode: 284, Score: -124.21376390228994  
Episode: 285, Score: -127.44845553614408  
Episode: 286, Score: -125.82824702221039  
Episode: 287, Score: -11.383789722760307  
Episode: 288, Score: -132.7717567474053  
Episode: 289, Score: -134.77534509411035  
Episode: 290, Score: -123.5835239301171  
Episode: 291, Score: -4.722876152646389  
Episode: 292, Score: -127.66064531224131  
Episode: 293, Score: -251.88893165444702  
Episode: 294, Score: -6.215592142435282  
Episode: 295, Score: -6.990223785859828  
Episode: 296, Score: -116.32498347353373  
Episode: 297, Score: -231.30028462618293  
Episode: 298, Score: -244.71382540762627  
Episode: 299, Score: -4.412290325294387

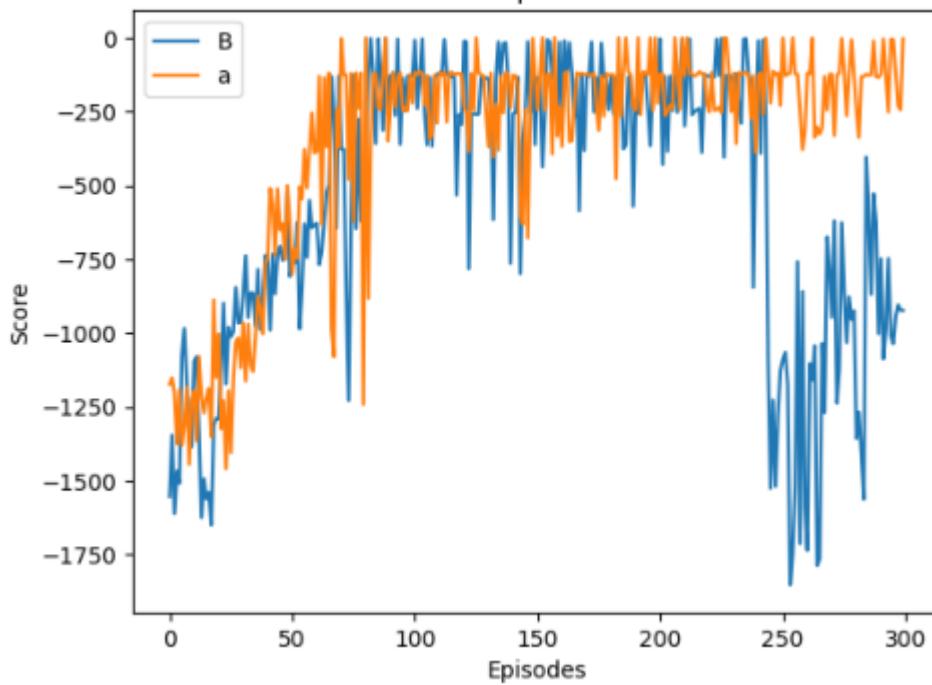


Comparison



```
In [66]: torch.save(improvedAgent.policy_net.state_dict(), "./rl_models/improved.pt")
showImage("./Img/RL/added_leaky_relu.png")
```

Comparison



```
In [42]: evaluate_model(baseAgent, 100)
```

Evaluating Episode: 1, Score: -1201.1995185104172  
Evaluating Episode: 2, Score: -1186.1421552471024  
Evaluating Episode: 3, Score: -1191.1385639932867  
Evaluating Episode: 4, Score: -1201.9092745921062  
Evaluating Episode: 5, Score: -1190.8787614738426  
Evaluating Episode: 6, Score: -804.5394741620818  
Evaluating Episode: 7, Score: -1221.1918615899008  
Evaluating Episode: 8, Score: -1188.9339394180201  
Evaluating Episode: 9, Score: -1190.5634314348347  
Evaluating Episode: 10, Score: -1195.2881979093152  
Evaluating Episode: 11, Score: -1183.562516205666  
Evaluating Episode: 12, Score: -1187.6975535956753  
Evaluating Episode: 13, Score: -1184.2673682234838  
Evaluating Episode: 14, Score: -1211.3206536406162  
Evaluating Episode: 15, Score: -1220.0047354615392  
Evaluating Episode: 16, Score: -1182.1586356337962  
Evaluating Episode: 17, Score: -1171.9233514512623  
Evaluating Episode: 18, Score: -1198.6910967963877  
Evaluating Episode: 19, Score: -1044.6449134294276  
Evaluating Episode: 20, Score: -1186.2231497246648  
Evaluating Episode: 21, Score: -1165.350319157556  
Evaluating Episode: 22, Score: -1214.8660451964302  
Evaluating Episode: 23, Score: -1265.5170396703302  
Evaluating Episode: 24, Score: -1192.5854123831673  
Evaluating Episode: 25, Score: -1202.773634212631  
Evaluating Episode: 26, Score: -1232.4998630479872  
Evaluating Episode: 27, Score: -1196.0146084746361  
Evaluating Episode: 28, Score: -1072.8646054749538  
Evaluating Episode: 29, Score: -1189.4727585027242  
Evaluating Episode: 30, Score: -1185.0048391622927  
Evaluating Episode: 31, Score: -1071.2540750658463  
Evaluating Episode: 32, Score: -1188.0772297880562  
Evaluating Episode: 33, Score: -1203.1911218866294  
Evaluating Episode: 34, Score: -1199.2363164156495  
Evaluating Episode: 35, Score: -1160.520865153369  
Evaluating Episode: 36, Score: -1198.7764980822778  
Evaluating Episode: 37, Score: -1198.9828514956753  
Evaluating Episode: 38, Score: -1202.2606634678023  
Evaluating Episode: 39, Score: -1225.7474905350525  
Evaluating Episode: 40, Score: -1264.9272437473496  
Evaluating Episode: 41, Score: -1031.4329590749276  
Evaluating Episode: 42, Score: -1205.2060152792358  
Evaluating Episode: 43, Score: -1198.503156882298  
Evaluating Episode: 44, Score: -1186.7498161770563  
Evaluating Episode: 45, Score: -1186.4301685308897  
Evaluating Episode: 46, Score: -1251.730120288701  
Evaluating Episode: 47, Score: -1185.39683395333  
Evaluating Episode: 48, Score: -1203.8824058550422  
Evaluating Episode: 49, Score: -1169.8611744191855  
Evaluating Episode: 50, Score: -1169.4389451270602  
Evaluating Episode: 51, Score: -1192.6206979243227  
Evaluating Episode: 52, Score: -1186.3888556977745  
Evaluating Episode: 53, Score: -1118.1701231353666  
Evaluating Episode: 54, Score: -1198.3592302485276  
Evaluating Episode: 55, Score: -1066.6269848566208  
Evaluating Episode: 56, Score: -1197.2899301625423  
Evaluating Episode: 57, Score: -1193.8854601383962  
Evaluating Episode: 58, Score: -1205.1850587698407  
Evaluating Episode: 59, Score: -1216.8977261431137  
Evaluating Episode: 60, Score: -1035.1458746363294  
Evaluating Episode: 61, Score: -1198.0365334557957  
Evaluating Episode: 62, Score: -1202.5109564168392  
Evaluating Episode: 63, Score: -1199.1453050499942  
Evaluating Episode: 64, Score: -1163.6989750770688  
Evaluating Episode: 65, Score: -1190.0511480341233  
Evaluating Episode: 66, Score: -1195.0355509926803

```
Evaluating Episode: 67, Score: -1138.9910056053477  
Evaluating Episode: 68, Score: -1184.8107800420948  
Evaluating Episode: 69, Score: -1196.8622752838355  
Evaluating Episode: 70, Score: -1188.6924355436524  
Evaluating Episode: 71, Score: -1228.785169981276  
Evaluating Episode: 72, Score: -1179.5078925182836  
Evaluating Episode: 73, Score: -1203.7859224312397  
Evaluating Episode: 74, Score: -1159.9889914826808  
Evaluating Episode: 75, Score: -1190.3306602586008  
Evaluating Episode: 76, Score: -1199.5205093494383  
Evaluating Episode: 77, Score: -1201.9147440380789  
Evaluating Episode: 78, Score: -1197.3006237457448  
Evaluating Episode: 79, Score: -1200.7739803602965  
Evaluating Episode: 80, Score: -1207.1747607429015  
Evaluating Episode: 81, Score: -1179.3359446705092  
Evaluating Episode: 82, Score: -1199.6298354556673  
Evaluating Episode: 83, Score: -1203.0424326424065  
Evaluating Episode: 84, Score: -1076.4392842217055  
Evaluating Episode: 85, Score: -1194.3917321489255  
Evaluating Episode: 86, Score: -1168.156246133869  
Evaluating Episode: 87, Score: -1177.3935645381155  
Evaluating Episode: 88, Score: -1200.4878167618326  
Evaluating Episode: 89, Score: -1142.6193236927968  
Evaluating Episode: 90, Score: -1203.138519873201  
Evaluating Episode: 91, Score: -1181.5622083238409  
Evaluating Episode: 92, Score: -1054.0099947898414  
Evaluating Episode: 93, Score: -1191.9106932326517  
Evaluating Episode: 94, Score: -1194.856840651913  
Evaluating Episode: 95, Score: -1236.1934098675747  
Evaluating Episode: 96, Score: -1199.6618537541433  
Evaluating Episode: 97, Score: -1200.8151886413714  
Evaluating Episode: 98, Score: -1208.7774377668663  
Evaluating Episode: 99, Score: -1148.2704851705016  
Evaluating Episode: 100, Score: -1196.3470533154957  
Final Score: -1179.553622567776
```

```
In [43]: evaluate_model(improvedAgent, 100)
```

Evaluating Episode: 1, Score: -125.19273663064577  
Evaluating Episode: 2, Score: -228.88074451907121  
Evaluating Episode: 3, Score: -1.5658796387360794  
Evaluating Episode: 4, Score: -120.38888903218249  
Evaluating Episode: 5, Score: -115.08704786142602  
Evaluating Episode: 6, Score: -121.00840725538347  
Evaluating Episode: 7, Score: -125.52391964159162  
Evaluating Episode: 8, Score: -1.7934736986394302  
Evaluating Episode: 9, Score: -128.36517999596995  
Evaluating Episode: 10, Score: -247.92646476893867  
Evaluating Episode: 11, Score: -125.27701220511095  
Evaluating Episode: 12, Score: -447.5055630717795  
Evaluating Episode: 13, Score: -232.62556224675754  
Evaluating Episode: 14, Score: -3.126555459554056  
Evaluating Episode: 15, Score: -238.71413995382366  
Evaluating Episode: 16, Score: -324.03139511078706  
Evaluating Episode: 17, Score: -125.68939970323304  
Evaluating Episode: 18, Score: -1.702359255051591  
Evaluating Episode: 19, Score: -122.30174980043112  
Evaluating Episode: 20, Score: -376.4314298540044  
Evaluating Episode: 21, Score: -125.4336267291072  
Evaluating Episode: 22, Score: -1.313470737973587  
Evaluating Episode: 23, Score: -128.48250026775116  
Evaluating Episode: 24, Score: -230.6797813335825  
Evaluating Episode: 25, Score: -126.40629431592993  
Evaluating Episode: 26, Score: -2.271593743046745  
Evaluating Episode: 27, Score: -121.08032906430928  
Evaluating Episode: 28, Score: -240.44475012162604  
Evaluating Episode: 29, Score: -122.05802882153588  
Evaluating Episode: 30, Score: -126.13918130952189  
Evaluating Episode: 31, Score: -232.97709770419266  
Evaluating Episode: 32, Score: -337.19343392958876  
Evaluating Episode: 33, Score: -241.37915739730627  
Evaluating Episode: 34, Score: -124.83117978628538  
Evaluating Episode: 35, Score: -123.5980864602523  
Evaluating Episode: 36, Score: -242.48297189678263  
Evaluating Episode: 37, Score: -121.03470155504932  
Evaluating Episode: 38, Score: -251.99264843326455  
Evaluating Episode: 39, Score: -242.6564923243504  
Evaluating Episode: 40, Score: -1.4744452448771213  
Evaluating Episode: 41, Score: -1.3588737976775755  
Evaluating Episode: 42, Score: -357.32031346112234  
Evaluating Episode: 43, Score: -127.8116474558748  
Evaluating Episode: 44, Score: -129.43028312182673  
Evaluating Episode: 45, Score: -353.910913909013  
Evaluating Episode: 46, Score: -116.8861650479959  
Evaluating Episode: 47, Score: -127.60662643754269  
Evaluating Episode: 48, Score: -125.45749454825895  
Evaluating Episode: 49, Score: -124.6954336550945  
Evaluating Episode: 50, Score: -236.69626114374896  
Evaluating Episode: 51, Score: -117.31213476516697  
Evaluating Episode: 52, Score: -236.7435956549321  
Evaluating Episode: 53, Score: -121.31590273950461  
Evaluating Episode: 54, Score: -250.45459547662554  
Evaluating Episode: 55, Score: -123.81409815023923  
Evaluating Episode: 56, Score: -126.75683219849678  
Evaluating Episode: 57, Score: -240.0613739057459  
Evaluating Episode: 58, Score: -355.2192907911392  
Evaluating Episode: 59, Score: -230.8042394252449  
Evaluating Episode: 60, Score: -345.9991394024025  
Evaluating Episode: 61, Score: -226.1005182215871  
Evaluating Episode: 62, Score: -123.94954591797435  
Evaluating Episode: 63, Score: -120.27992601335474  
Evaluating Episode: 64, Score: -231.34941849686163  
Evaluating Episode: 65, Score: -2.044680518888446  
Evaluating Episode: 66, Score: -121.8612252758446

```
Evaluating Episode: 67, Score: -118.47572212521911
Evaluating Episode: 68, Score: -121.2846974724934
Evaluating Episode: 69, Score: -121.4301066845396
Evaluating Episode: 70, Score: -316.9097987460727
Evaluating Episode: 71, Score: -126.40999407616464
Evaluating Episode: 72, Score: -2.463241279423309
Evaluating Episode: 73, Score: -285.7966936618585
Evaluating Episode: 74, Score: -122.11123749206247
Evaluating Episode: 75, Score: -241.25164901994714
Evaluating Episode: 76, Score: -3.300934392278764
Evaluating Episode: 77, Score: -120.41514841441192
Evaluating Episode: 78, Score: -126.37744293952056
Evaluating Episode: 79, Score: -1.5083511896029524
Evaluating Episode: 80, Score: -123.79063528472724
Evaluating Episode: 81, Score: -239.13350790255035
Evaluating Episode: 82, Score: -117.80986542090113
Evaluating Episode: 83, Score: -126.26347024552327
Evaluating Episode: 84, Score: -234.71725804027298
Evaluating Episode: 85, Score: -122.82980447481786
Evaluating Episode: 86, Score: -1.4178313351031209
Evaluating Episode: 87, Score: -225.82531118290362
Evaluating Episode: 88, Score: -128.28554550905457
Evaluating Episode: 89, Score: -117.62580937835203
Evaluating Episode: 90, Score: -125.9508652249148
Evaluating Episode: 91, Score: -239.58636849856921
Evaluating Episode: 92, Score: -118.03860242300429
Evaluating Episode: 93, Score: -238.97263216603804
Evaluating Episode: 94, Score: -115.3281884804444
Evaluating Episode: 95, Score: -1.5305842457641325
Evaluating Episode: 96, Score: -246.29451172421435
Evaluating Episode: 97, Score: -126.82917885684667
Evaluating Episode: 98, Score: -123.11075482791087
Evaluating Episode: 99, Score: -127.20183332803614
Evaluating Episode: 100, Score: -115.36759250036938
Final Score: -157.55953378955527
```

Based on the graph, I find the Leaky\_Relu to have improved the model slightly. At the beginning, I feel that the mode improved faster. but nevertheless its a slight improvement.

## Training Stabilization

### Greedy Epsilon

During training, incorporating the Greedy Epsilon strategy enhances the stability of your Deep Q-Network (DQN) by managing the trade-off between exploration and exploitation. I hope this will stabilise the intial part of the training process.

- **What:** The Greedy Epsilon strategy allows the agent to choose between exploiting the best-known action with high probability and exploring new actions with a controlled probability.
- **Why:** By dynamically adjusting the exploration rate, the model avoids premature convergence to suboptimal policies. This ensures a more balanced exploration of the state-action space, preventing the DQN from getting stuck in local minima and facilitating a more stable learning process.

---

**Algorithm 1:** Epsilon-Greedy Q-Learning Algorithm

---

**Data:**  $\alpha$ : learning rate,  $\gamma$ : discount factor,  $\epsilon$ : a small number  
**Result:** A Q-table containing  $Q(S,A)$  pairs defining estimated optimal policy  $\pi^*$

```
/* Initialization */  
Initialize Q(s,a) arbitrarily, except Q(terminal, .);  
Q(terminal, .) ← 0;  
/* For each step in each episode, we calculate the Q-value and update the Q-table */  
for each episode do  
    /* Initialize state S, usually by resetting the environment */  
    Initialize state S;  
    for each step in episode do  
        do  
            /* Choose action A from S using epsilon-greedy policy derived from Q */  
            A ← SELECT-ACTION(Q, S,  $\epsilon$ );  
            Take action A, then observe reward R and next state S';  
            Q(S, A) ← Q(S, A) +  $\alpha$  [ R +  $\gamma \max_a Q(S', a)$  - Q(S, A)];  
            S ← S';  
        while S is not terminal;  
    end  
end
```

---

## Gradient Clipping

By controlling how much the model can change, I hope to mitigate "catastrophic forgetting" among the DQN.

- **What:** Gradient clipping involves limiting the magnitude of gradients during backpropagation to prevent extreme updates to the model's parameters.
- **Why:** By constraining the gradients, the training process becomes more robust, preventing numerical instability issues associated with exploding gradients. This leads to smoother convergence during optimization and enhances the stability of your DQN model.

To implement these two, we would need to edit the DQNAgent and the training loop.

```
In [55]: class DQNAgentEpsGreedy:  
    def __init__(self):  
        self.steps = 0  
        self.memory = Memory(10000)  
  
        # Greedy Epsilon Hyperparams  
        self.epsilon = 1.0  
        self.epsilon_decay = 0.98  
        self.epsilon_min = 0.001  
  
        self.batch_size = 256  
        self.adamlr = 0.001  
        self.gamma = 0.98  
  
        self.policy_net = MLP(3, 9, self.adamlr)  
  
    def action(self, state):  
        sample = np.random.rand()  
  
        if sample > self.epsilon:  
            with torch.no_grad():
```

```

        choice = float(torch.argmax(self.policy_net(state)).numpy())
        action = (choice-4)/2

    return choice, action

def optimize(self):
    if len(self.memory) < self.batch_size:
        return

    batch = self.memory.sample(self.batch_size)
    state, action, reward, next_state = batch
    action = action.type(torch.int64)

    # target
    with torch.no_grad(): # parameters from previous iteration are held fixed
        yi = reward + self.gamma * self.policy_net(next_state).max(1)[0].unsqueeze(0)

    # Q(phi_i, a_i, theta)
    current_state_action = self.policy_net(state).gather(1,action)

    # Back_propagation
    loss = (yi - current_state_action) ** 2
    self.policy_net.optimizer.zero_grad()
    loss.mean().backward()
    nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
    self.policy_net.optimizer.step()

```

In [56]:

```

def train_model_eps(DQNAgentX, episodes):
    agent = DQNAgentX()
    env = gym.make('Pendulum-v1')
    score_list = []
    for episode in range(episodes):
        print(f"Episode: {episode}", end="")
        state, _ = env.reset()
        state = torch.tensor(state)
        done = False
        score = 0
        while not done:
            choice, action = agent.action(state)

            next_state, reward, terminated, truncated, _ = env.step([action])

            score+=reward
            done = terminated or truncated

            if terminated:
                next_state = None
            else:
                next_state = torch.tensor(next_state,dtype=torch.float32)

            agent.memory.push((state, choice, reward, next_state ))
            state = next_state

            if len(agent.memory) > 1000:
                agent.optimize()

        print(f" Score: {score}, Threshold: {agent.epsilon}")
        score_list.append(score)

        if agent.epsilon > agent.epsilon_min:
            agent.epsilon *= agent.epsilon_decay
    env.close()

```

```
plt.plot(score_list)
plt.show()
```

```
return agent, score_list
```

```
In [57]: epsAgent, eps_score = train_model_eps(DQNAgentEpsGreedy, 300)
plot_tuples([(improved_score,"base"),(eps_score,"eps")],"Compare Episilon and Base")
```

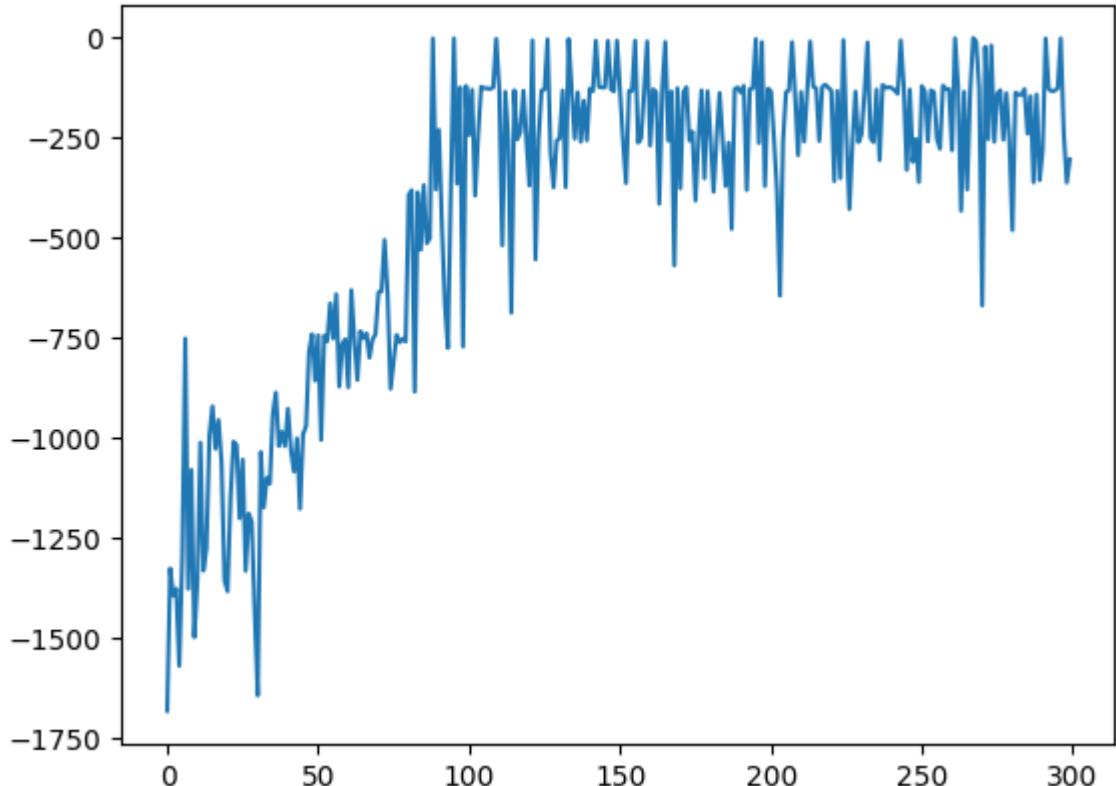
Episode: 0, Score: -1681.5943774854052, Threshold: 1.0  
Episode: 1, Score: -1326.1277777174917, Threshold: 0.98  
Episode: 2, Score: -1393.4720205026888, Threshold: 0.9603999999999999  
Episode: 3, Score: -1376.5442862355353, Threshold: 0.9411919999999999  
Episode: 4, Score: -1569.2645259349085, Threshold: 0.9223681599999999  
Episode: 5, Score: -1297.1408043712051, Threshold: 0.9039207967999998  
Episode: 6, Score: -751.728848818093, Threshold: 0.8858423808639998  
Episode: 7, Score: -1377.0481414803924, Threshold: 0.8681255332467198  
Episode: 8, Score: -1079.8173487575223, Threshold: 0.8507630225817854  
Episode: 9, Score: -1497.5934056795275, Threshold: 0.8337477621301497  
Episode: 10, Score: -1356.8526517111075, Threshold: 0.8170728068875467  
Episode: 11, Score: -1011.8216520388379, Threshold: 0.8007313507497957  
Episode: 12, Score: -1331.3294240433318, Threshold: 0.7847167237347998  
Episode: 13, Score: -1276.3049858017848, Threshold: 0.7690223892601038  
Episode: 14, Score: -992.0721316297617, Threshold: 0.7536419414749017  
Episode: 15, Score: -921.0875938136455, Threshold: 0.7385691026454037  
Episode: 16, Score: -1027.1131601681848, Threshold: 0.7237977205924956  
Episode: 17, Score: -955.2554768706651, Threshold: 0.7093217661806457  
Episode: 18, Score: -1060.6563501123337, Threshold: 0.6951353308570327  
Episode: 19, Score: -1356.1533904643275, Threshold: 0.6812326242398921  
Episode: 20, Score: -1382.775789784248, Threshold: 0.6676079717550942  
Episode: 21, Score: -1146.045843847422, Threshold: 0.6542558123199923  
Episode: 22, Score: -1009.362839003556, Threshold: 0.6411706960735924  
Episode: 23, Score: -1018.6688270269418, Threshold: 0.6283472821521205  
Episode: 24, Score: -1200.9895671156305, Threshold: 0.6157803365090782  
Episode: 25, Score: -1054.30875104957, Threshold: 0.6034647297788965  
Episode: 26, Score: -1332.167242466003, Threshold: 0.5913954351833186  
Episode: 27, Score: -1189.0430564124056, Threshold: 0.5795675264796523  
Episode: 28, Score: -1209.4474154553452, Threshold: 0.5679761759500592  
Episode: 29, Score: -1435.285689611141, Threshold: 0.5566166524310581  
Episode: 30, Score: -1643.1951805616277, Threshold: 0.5454843193824369  
Episode: 31, Score: -1035.1773779180514, Threshold: 0.5345746329947881  
Episode: 32, Score: -1173.5728769032442, Threshold: 0.5238831403348924  
Episode: 33, Score: -1100.167523427386, Threshold: 0.5134054775281945  
Episode: 34, Score: -1114.0421727965368, Threshold: 0.5031373679776306  
Episode: 35, Score: -939.8631123732987, Threshold: 0.493074620618078  
Episode: 36, Score: -886.7659316579415, Threshold: 0.48321312820571644  
Episode: 37, Score: -1020.2944644616096, Threshold: 0.4735488656416021  
Episode: 38, Score: -984.8843569454353, Threshold: 0.46407788832877006  
Episode: 39, Score: -1020.4621246082679, Threshold: 0.45479633056219465  
Episode: 40, Score: -927.227937534528, Threshold: 0.44570040395095073  
Episode: 41, Score: -1030.7060294797031, Threshold: 0.4367863958719317  
Episode: 42, Score: -1084.1634949597917, Threshold: 0.42805066795449304  
Episode: 43, Score: -1001.317348141445, Threshold: 0.41948965459540316  
Episode: 44, Score: -1176.8877279611277, Threshold: 0.4110998615034951  
Episode: 45, Score: -990.9884203119846, Threshold: 0.4028778642734252  
Episode: 46, Score: -970.5547169444015, Threshold: 0.39482030698795667  
Episode: 47, Score: -784.4782585897507, Threshold: 0.38692390084819756  
Episode: 48, Score: -741.2499015053831, Threshold: 0.3791854228312336  
Episode: 49, Score: -857.2127479851099, Threshold: 0.37160171437460887  
Episode: 50, Score: -743.8700406187639, Threshold: 0.3641696800871167  
Episode: 51, Score: -1005.0615634962326, Threshold: 0.35688628648537435  
Episode: 52, Score: -744.6545776904196, Threshold: 0.34974856075566685  
Episode: 53, Score: -758.7754428203099, Threshold: 0.3427535895405535  
Episode: 54, Score: -664.057500921372, Threshold: 0.33589851774974244  
Episode: 55, Score: -751.8490292973535, Threshold: 0.3291805473947476  
Episode: 56, Score: -641.1822679323822, Threshold: 0.32259693644685267  
Episode: 57, Score: -871.8267202426026, Threshold: 0.3161449977179156  
Episode: 58, Score: -765.6028516033703, Threshold: 0.3098220977635573  
Episode: 59, Score: -753.5225900584668, Threshold: 0.30362565580828615  
Episode: 60, Score: -873.5198557377, Threshold: 0.2975531426921204  
Episode: 61, Score: -631.202341454779, Threshold: 0.291602079838278  
Episode: 62, Score: -755.454734508629, Threshold: 0.2857700382415124  
Episode: 63, Score: -855.3640470142217, Threshold: 0.2800546374766822  
Episode: 64, Score: -733.6406501981048, Threshold: 0.27445354472714856  
Episode: 65, Score: -750.9669627369648, Threshold: 0.2689644738326056

Episode: 66, Score: -739.0062568177694, Threshold: 0.26358518435595346  
Episode: 67, Score: -799.5152984916756, Threshold: 0.25831348066883436  
Episode: 68, Score: -754.6220930559516, Threshold: 0.2531472110554577  
Episode: 69, Score: -740.9741286944211, Threshold: 0.24808426683434853  
Episode: 70, Score: -635.9968507498841, Threshold: 0.24312258149766156  
Episode: 71, Score: -633.8322851634281, Threshold: 0.2382601298677083  
Episode: 72, Score: -505.6110917649418, Threshold: 0.23349492727035415  
Episode: 73, Score: -637.7304897125906, Threshold: 0.22882502872494706  
Episode: 74, Score: -877.4005652512267, Threshold: 0.22424852815044813  
Episode: 75, Score: -808.0637648428379, Threshold: 0.21976355758743915  
Episode: 76, Score: -742.8065389062266, Threshold: 0.21536828643569036  
Episode: 77, Score: -761.3582477832698, Threshold: 0.21106092070697655  
Episode: 78, Score: -752.26647466192, Threshold: 0.20683970229283702  
Episode: 79, Score: -759.3707024596908, Threshold: 0.20270290824698028  
Episode: 80, Score: -392.97646701216667, Threshold: 0.19864885008204067  
Episode: 81, Score: -381.4281496544306, Threshold: 0.19467587308039985  
Episode: 82, Score: -884.616584678361, Threshold: 0.19078235561879187  
Episode: 83, Score: -387.0279001763854, Threshold: 0.18696670850641603  
Episode: 84, Score: -530.0282838546775, Threshold: 0.18322737433628772  
Episode: 85, Score: -368.9237435280614, Threshold: 0.17956282684956196  
Episode: 86, Score: -513.963957249899, Threshold: 0.1759715703125707  
Episode: 87, Score: -501.77912415832014, Threshold: 0.1724521389063193  
Episode: 88, Score: -2.2610423818947045, Threshold: 0.16900309612819292  
Episode: 89, Score: -379.67840925353295, Threshold: 0.16562303420562907  
Episode: 90, Score: -230.5792250393103, Threshold: 0.16231057352151648  
Episode: 91, Score: -468.41195467449717, Threshold: 0.15906436205108615  
Episode: 92, Score: -666.6172472514029, Threshold: 0.15588307481006441  
Episode: 93, Score: -775.4545695968687, Threshold: 0.15276541331386312  
Episode: 94, Score: -364.40437470037335, Threshold: 0.14971010504758586  
Episode: 95, Score: -2.161498807395704, Threshold: 0.14671590294663414  
Episode: 96, Score: -364.93699375270495, Threshold: 0.14378158488770146  
Episode: 97, Score: -124.78885918244748, Threshold: 0.14090595318994742  
Episode: 98, Score: -771.9741005854123, Threshold: 0.13808783412614847  
Episode: 99, Score: -121.28422941718506, Threshold: 0.13532607744362551  
Episode: 100, Score: -243.57710780233933, Threshold: 0.132619555894753  
Episode: 101, Score: -130.57132553594403, Threshold: 0.12996716477685794  
Episode: 102, Score: -395.04018387493494, Threshold: 0.12736782148132078  
Episode: 103, Score: -245.0078528518176, Threshold: 0.12482046505169436  
Episode: 104, Score: -123.41133578303032, Threshold: 0.12232405575066048  
Episode: 105, Score: -126.33391203700256, Threshold: 0.11987757463564727  
Episode: 106, Score: -127.42895941712044, Threshold: 0.11748002314293432  
Episode: 107, Score: -128.40848935477834, Threshold: 0.11513042268007563  
Episode: 108, Score: -124.67876817121247, Threshold: 0.11282781422647412  
Episode: 109, Score: -3.5005167858570005, Threshold: 0.11057125794194463  
Episode: 110, Score: -128.93911799634142, Threshold: 0.10835983278310574  
Episode: 111, Score: -518.9822603802734, Threshold: 0.10619263612744362  
Episode: 112, Score: -135.83420349683817, Threshold: 0.10406878340489474  
Episode: 113, Score: -257.6115438845531, Threshold: 0.10198740773679685  
Episode: 114, Score: -687.386071810616, Threshold: 0.09994765958206091  
Episode: 115, Score: -132.28184829325892, Threshold: 0.0979487063904197  
Episode: 116, Score: -254.21820143801486, Threshold: 0.0959897322626113  
Episode: 117, Score: -235.30005218781508, Threshold: 0.09406993761735907  
Episode: 118, Score: -133.10003930949603, Threshold: 0.0921885388650119  
Episode: 119, Score: -257.1173225453407, Threshold: 0.09034476808771166  
Episode: 120, Score: -369.6318212175632, Threshold: 0.08853787272595742  
Episode: 121, Score: -7.375531255962358, Threshold: 0.08676711527143827  
Episode: 122, Score: -554.3613514112941, Threshold: 0.0850317729660095  
Episode: 123, Score: -255.05753613455394, Threshold: 0.08333113750668932  
Episode: 124, Score: -132.92244958067246, Threshold: 0.08166451475655553  
Episode: 125, Score: -129.80291637776517, Threshold: 0.08003122446142442  
Episode: 126, Score: -4.884771168320977, Threshold: 0.07843059997219594  
Episode: 127, Score: -287.5297843086242, Threshold: 0.07686198797275202  
Episode: 128, Score: -374.1882560883133, Threshold: 0.07532474821329699  
Episode: 129, Score: -258.0830691151863, Threshold: 0.07381825324903105  
Episode: 130, Score: -252.05383494212364, Threshold: 0.07234188818405042  
Episode: 131, Score: -132.37830325710263, Threshold: 0.07089505042036941

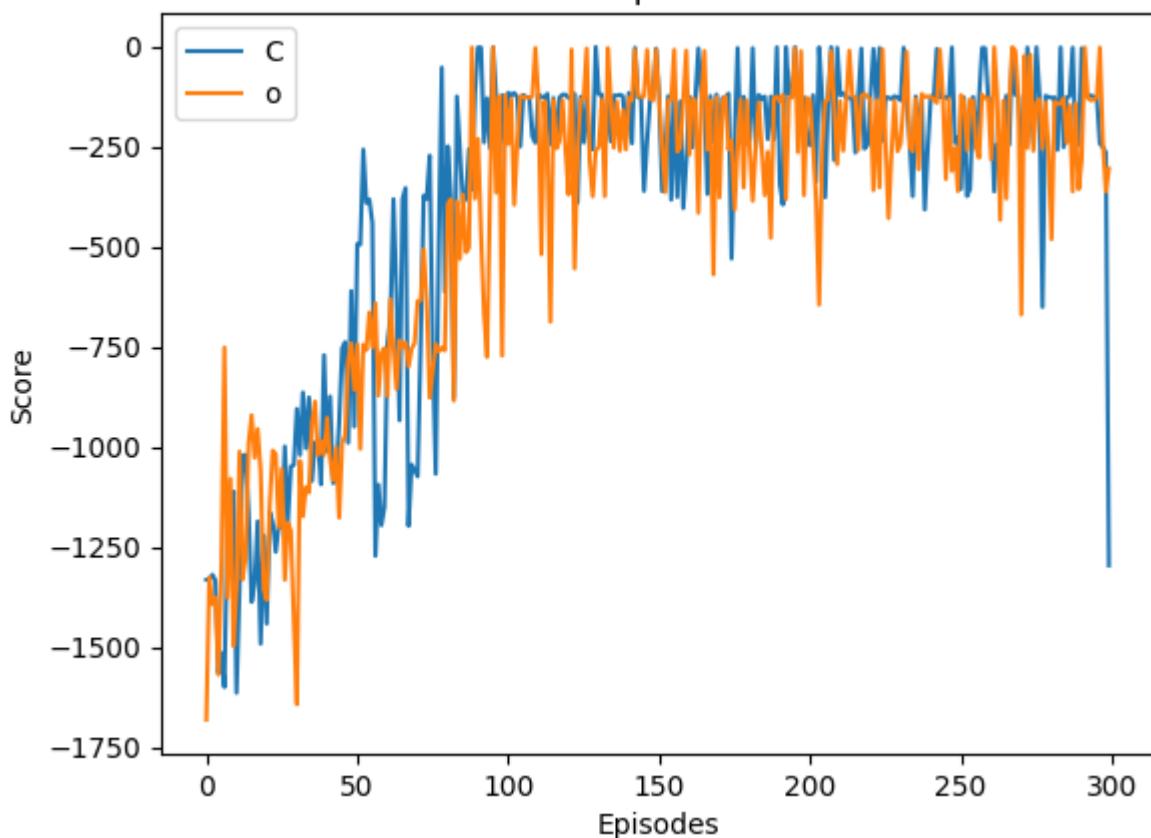
Episode: 132, Score: -373.8703662868735, Threshold: 0.06947714941196202  
Episode: 133, Score: -3.7888366112092187, Threshold: 0.06808760642372277  
Episode: 134, Score: -133.09977381171583, Threshold: 0.06672585429524831  
Episode: 135, Score: -252.76279751858425, Threshold: 0.06539133720934334  
Episode: 136, Score: -137.25876080456132, Threshold: 0.06408351046515648  
Episode: 137, Score: -260.49911865572517, Threshold: 0.06280184025585335  
Episode: 138, Score: -158.0232023794963, Threshold: 0.061545803450736285  
Episode: 139, Score: -256.9855437579881, Threshold: 0.060314887381721555  
Episode: 140, Score: -128.0095000518157, Threshold: 0.05910858963408712  
Episode: 141, Score: -134.3575978095787, Threshold: 0.05792641784140538  
Episode: 142, Score: -8.047797391652548, Threshold: 0.05676788948457727  
Episode: 143, Score: -123.1937657485461, Threshold: 0.055632531694885724  
Episode: 144, Score: -125.07297771706264, Threshold: 0.054519881060988006  
Episode: 145, Score: -124.66199680463443, Threshold: 0.05342948343976824  
Episode: 146, Score: -8.300106740954625, Threshold: 0.052360893770972874  
Episode: 147, Score: -129.7516508326815, Threshold: 0.05131367589555342  
Episode: 148, Score: -134.3695694481085, Threshold: 0.05028740237764235  
Episode: 149, Score: -7.819905872666513, Threshold: 0.0492816543300895  
Episode: 150, Score: -129.54308563291585, Threshold: 0.04829602124348771  
Episode: 151, Score: -256.89869092847886, Threshold: 0.04733010081861796  
Episode: 152, Score: -363.6188022182613, Threshold: 0.0463834988022456  
Episode: 153, Score: -133.41206881879722, Threshold: 0.04545582882620069  
Episode: 154, Score: -133.35359834903892, Threshold: 0.04454671224967667  
Episode: 155, Score: -7.939581279243796, Threshold: 0.043655778004683135  
Episode: 156, Score: -262.5821911031511, Threshold: 0.04278266244458947  
Episode: 157, Score: -251.7317822874901, Threshold: 0.04192700919569768  
Episode: 158, Score: -127.60896631622714, Threshold: 0.04108846901178373  
Episode: 159, Score: -10.137892988996786, Threshold: 0.04026669963154805  
Episode: 160, Score: -270.0396145851201, Threshold: 0.03946136563891709  
Episode: 161, Score: -129.6724957030623, Threshold: 0.03867213832613875  
Episode: 162, Score: -136.6578423080876, Threshold: 0.037898695559615975  
Episode: 163, Score: -415.38478336998486, Threshold: 0.03714072164842366  
Episode: 164, Score: -128.99389637517947, Threshold: 0.03639790721545518  
Episode: 165, Score: -11.315629690741092, Threshold: 0.035669949071146075  
Episode: 166, Score: -257.77036743074945, Threshold: 0.034956550089723155  
Episode: 167, Score: -134.80816959065208, Threshold: 0.03425741908792869  
Episode: 168, Score: -569.2809229173293, Threshold: 0.03357227070617012  
Episode: 169, Score: -126.61803435271182, Threshold: 0.032900825292046715  
Episode: 170, Score: -376.6998621718786, Threshold: 0.03224280878620578  
Episode: 171, Score: -137.12417129548714, Threshold: 0.031597952610481664  
Episode: 172, Score: -124.31011541812919, Threshold: 0.03096599355827203  
Episode: 173, Score: -257.32205205566885, Threshold: 0.030346673687106588  
Episode: 174, Score: -235.37857923116246, Threshold: 0.029739740213364455  
Episode: 175, Score: -407.46180907146663, Threshold: 0.029144945409097164  
Episode: 176, Score: -248.93753743895266, Threshold: 0.02856204650091522  
Episode: 177, Score: -132.17840710154357, Threshold: 0.027990805570896914  
Episode: 178, Score: -351.94071156032754, Threshold: 0.027430989459478975  
Episode: 179, Score: -134.43125019387588, Threshold: 0.026882369670289397  
Episode: 180, Score: -252.01214342395912, Threshold: 0.026344722276883607  
Episode: 181, Score: -385.30838590655907, Threshold: 0.025817827831345935  
Episode: 182, Score: -252.09905468991204, Threshold: 0.025301471274719014  
Episode: 183, Score: -138.8196489858593, Threshold: 0.024795441849224633  
Episode: 184, Score: -257.8931478924562, Threshold: 0.02429953301224014  
Episode: 185, Score: -371.1583962820556, Threshold: 0.02381354235199534  
Episode: 186, Score: -263.0706362190519, Threshold: 0.02333727150495543  
Episode: 187, Score: -478.03906946317204, Threshold: 0.02287052607485632  
Episode: 188, Score: -129.1424223828612, Threshold: 0.022413115553359196  
Episode: 189, Score: -124.91018916519064, Threshold: 0.021964853242292012  
Episode: 190, Score: -137.64877630678683, Threshold: 0.02152555617744617  
Episode: 191, Score: -120.72612565353447, Threshold: 0.021095045053897248  
Episode: 192, Score: -381.11704907114466, Threshold: 0.020673144152819303  
Episode: 193, Score: -128.94925130735197, Threshold: 0.02025968126976292  
Episode: 194, Score: -127.13007994722967, Threshold: 0.01985448764436766  
Episode: 195, Score: -3.9377178621710303, Threshold: 0.019457397891480308  
Episode: 196, Score: -263.7132839790126, Threshold: 0.0190682499336507  
Episode: 197, Score: -11.604248684608043, Threshold: 0.018686884934977686

Episode: 198, Score: -371.2685045255932, Threshold: 0.01831314723627813  
Episode: 199, Score: -127.66278390603208, Threshold: 0.017946884291552567  
Episode: 200, Score: -134.31186611019479, Threshold: 0.017587946605721515  
Episode: 201, Score: -251.2884447814736, Threshold: 0.017236187673607083  
Episode: 202, Score: -382.82690928126823, Threshold: 0.016891463920134943  
Episode: 203, Score: -644.8127879220237, Threshold: 0.016553634641732245  
Episode: 204, Score: -259.4352606633373, Threshold: 0.0162225619488976  
Episode: 205, Score: -136.40687518267708, Threshold: 0.015898110709919647  
Episode: 206, Score: -131.99942959021215, Threshold: 0.015580148495721254  
Episode: 207, Score: -11.538949696193324, Threshold: 0.015268545525806828  
Episode: 208, Score: -143.5553201823254, Threshold: 0.014963174615290691  
Episode: 209, Score: -293.96708538559176, Threshold: 0.014663911122984877  
Episode: 210, Score: -136.3203790748211, Threshold: 0.01437063290052518  
Episode: 211, Score: -260.13066288296926, Threshold: 0.014083220242514675  
Episode: 212, Score: -138.73597066340585, Threshold: 0.013801555837664382  
Episode: 213, Score: -10.192325324449431, Threshold: 0.013525524720911095  
Episode: 214, Score: -122.67278853441428, Threshold: 0.013255014226492874  
Episode: 215, Score: -127.0908623856239, Threshold: 0.012989913941963017  
Episode: 216, Score: -258.74165033553754, Threshold: 0.012730115663123756  
Episode: 217, Score: -123.31996505811361, Threshold: 0.012475513349861281  
Episode: 218, Score: -117.94267073804262, Threshold: 0.012226003082864055  
Episode: 219, Score: -125.56042092596947, Threshold: 0.011981483021206774  
Episode: 220, Score: -132.74409862622687, Threshold: 0.011741853360782638  
Episode: 221, Score: -358.91955022754956, Threshold: 0.011507016293566985  
Episode: 222, Score: -133.7863123422791, Threshold: 0.011276875967695645  
Episode: 223, Score: -351.66610313960643, Threshold: 0.011051338448341731  
Episode: 224, Score: -6.48253415218795, Threshold: 0.010830311679374895  
Episode: 225, Score: -237.91663988277836, Threshold: 0.010613705445787398  
Episode: 226, Score: -428.58606042798203, Threshold: 0.01040143133687165  
Episode: 227, Score: -262.52167724422867, Threshold: 0.010193402710134216  
Episode: 228, Score: -135.22934983595007, Threshold: 0.009989534655931531  
Episode: 229, Score: -261.59942095775597, Threshold: 0.0097897439628129  
Episode: 230, Score: -242.43656225190895, Threshold: 0.009593949083556642  
Episode: 231, Score: -139.21423376603516, Threshold: 0.009402070101885509  
Episode: 232, Score: -12.693623685863397, Threshold: 0.009214028699847799  
Episode: 233, Score: -250.34814602020248, Threshold: 0.009029748125850843  
Episode: 234, Score: -261.5759637708384, Threshold: 0.008849153163333826  
Episode: 235, Score: -130.10435887621517, Threshold: 0.008672170100067149  
Episode: 236, Score: -306.52964867068465, Threshold: 0.008498726698065806  
Episode: 237, Score: -118.4479482097713, Threshold: 0.00832875216410449  
Episode: 238, Score: -125.98539614133942, Threshold: 0.0081621771208224  
Episode: 239, Score: -124.23256805925541, Threshold: 0.007998933578405951  
Episode: 240, Score: -125.87918341396976, Threshold: 0.007838954906837833  
Episode: 241, Score: -132.18358235011075, Threshold: 0.007682175808701076  
Episode: 242, Score: -140.66670739401258, Threshold: 0.007528532292527055  
Episode: 243, Score: -7.355045340992545, Threshold: 0.007377961646676513  
Episode: 244, Score: -121.02630667083687, Threshold: 0.007230402413742983  
Episode: 245, Score: -330.952653293619, Threshold: 0.007085794365468123  
Episode: 246, Score: -130.21993367191368, Threshold: 0.00694407847815876  
Episode: 247, Score: -310.62025825737294, Threshold: 0.0068051969085955845  
Episode: 248, Score: -253.51320359471745, Threshold: 0.006669092970423673  
Episode: 249, Score: -360.992269695945, Threshold: 0.00653571111015199  
Episode: 250, Score: -121.25392577800132, Threshold: 0.006404996888794895  
Episode: 251, Score: -135.23157827906482, Threshold: 0.006276896951018997  
Episode: 252, Score: -259.6302079872192, Threshold: 0.006151359011998617  
Episode: 253, Score: -132.05833576614154, Threshold: 0.006028331831758644  
Episode: 254, Score: -136.65174598240765, Threshold: 0.005907765195123471  
Episode: 255, Score: -253.0944467939218, Threshold: 0.005789609891221002  
Episode: 256, Score: -278.13729176998777, Threshold: 0.005673817693396582  
Episode: 257, Score: -120.31955858653944, Threshold: 0.00556034133952865  
Episode: 258, Score: -129.08421754792687, Threshold: 0.005449134512738076  
Episode: 259, Score: -129.25552128835488, Threshold: 0.0053401518224833145  
Episode: 260, Score: -281.8292017719162, Threshold: 0.005233348786033648  
Episode: 261, Score: -1.9045282979045393, Threshold: 0.005128681810312975  
Episode: 262, Score: -123.67365018078323, Threshold: 0.005026108174106716  
Episode: 263, Score: -432.95042136800544, Threshold: 0.004925586010624581

Episode: 264, Score: -135.52895428904876, Threshold: 0.004827074290412089  
Episode: 265, Score: -379.7434470552952, Threshold: 0.004730532804603847  
Episode: 266, Score: -136.28840313201988, Threshold: 0.00463592214851177  
Episode: 267, Score: -1.442788264691489, Threshold: 0.004543203705541535  
Episode: 268, Score: -10.732422538107258, Threshold: 0.004452339631430704  
Episode: 269, Score: -128.94873008122255, Threshold: 0.00436329283880209  
Episode: 270, Score: -669.3724956298362, Threshold: 0.004276026982026048  
Episode: 271, Score: -23.451604159817684, Threshold: 0.004190506442385527  
Episode: 272, Score: -253.67007096509195, Threshold: 0.004106696313537816  
Episode: 273, Score: -20.195699702492764, Threshold: 0.00402456238726706  
Episode: 274, Score: -260.6594085882816, Threshold: 0.003944071139521719  
Episode: 275, Score: -140.8661253285236, Threshold: 0.0038651897167312848  
Episode: 276, Score: -131.80347953338259, Threshold: 0.003787885922396659  
Episode: 277, Score: -255.49304828097195, Threshold: 0.003712128203948726  
Episode: 278, Score: -139.36285331151288, Threshold: 0.0036378856398697514  
Episode: 279, Score: -256.713321283043, Threshold: 0.0035651279270723563  
Episode: 280, Score: -481.32592946437705, Threshold: 0.0034938253685309092  
Episode: 281, Score: -137.7657487430093, Threshold: 0.003423948861160291  
Episode: 282, Score: -143.18463617344705, Threshold: 0.0033554698839370852  
Episode: 283, Score: -142.7787322880958, Threshold: 0.0032883604862583436  
Episode: 284, Score: -129.13163310026283, Threshold: 0.0032225932765331767  
Episode: 285, Score: -240.57877616652547, Threshold: 0.003158141411002513  
Episode: 286, Score: -146.6459047000545, Threshold: 0.003094978582782463  
Episode: 287, Score: -361.79107960514506, Threshold: 0.0030330790111268136  
Episode: 288, Score: -142.4231994834624, Threshold: 0.0029724174309042773  
Episode: 289, Score: -356.23803708245015, Threshold: 0.002912969082286192  
Episode: 290, Score: -282.0832628007327, Threshold: 0.002854709700640468  
Episode: 291, Score: -2.524217343673981, Threshold: 0.0027976155066276585  
Episode: 292, Score: -128.46392629162605, Threshold: 0.0027416631964951054  
Episode: 293, Score: -134.49783390016083, Threshold: 0.0026868299325652034  
Episode: 294, Score: -132.65311903059438, Threshold: 0.002633093333913899  
Episode: 295, Score: -127.06197250986791, Threshold: 0.002580431467235621  
Episode: 296, Score: -2.432722492336545, Threshold: 0.0025288228378909086  
Episode: 297, Score: -237.63707345727414, Threshold: 0.0024782463811330903  
Episode: 298, Score: -362.35324122738143, Threshold: 0.0024286814535104282  
Episode: 299, Score: -304.67093027832107, Threshold: 0.0023801078244402194

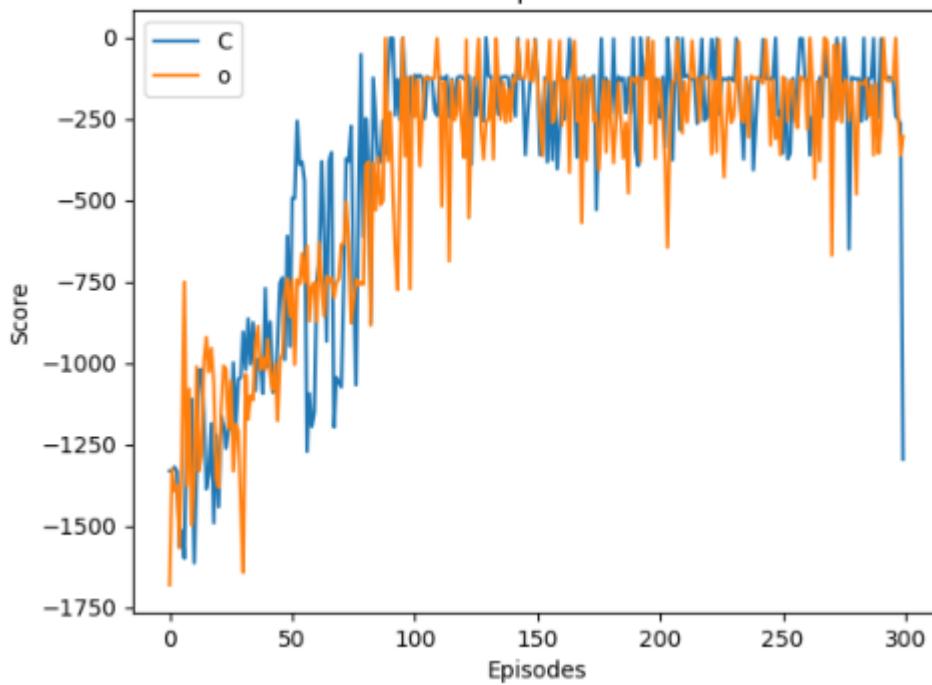


Comparison



```
In [67]: torch.save(epsAgent.policy_net.state_dict(), "./rl_models/epsModel.pt")
showImage("./Img/RL/eps_clipping.png")
```

Comparison



```
In [60]: evaluate_model(epsAgent, 100)
```

Evaluating Episode: 1, Score: -6.52191146639224  
Evaluating Episode: 2, Score: -428.82035319323376  
Evaluating Episode: 3, Score: -378.7240118946098  
Evaluating Episode: 4, Score: -127.04476226843582  
Evaluating Episode: 5, Score: -247.83476393311108  
Evaluating Episode: 6, Score: -131.83540507645583  
Evaluating Episode: 7, Score: -133.3815656584825  
Evaluating Episode: 8, Score: -244.2742441122501  
Evaluating Episode: 9, Score: -132.69417397307268  
Evaluating Episode: 10, Score: -376.0878705911997  
Evaluating Episode: 11, Score: -139.96788632523175  
Evaluating Episode: 12, Score: -6.192038872005059  
Evaluating Episode: 13, Score: -359.38003167871614  
Evaluating Episode: 14, Score: -6.6954756107901625  
Evaluating Episode: 15, Score: -250.18201890373277  
Evaluating Episode: 16, Score: -145.18148576589155  
Evaluating Episode: 17, Score: -335.42856279091774  
Evaluating Episode: 18, Score: -364.0482466542314  
Evaluating Episode: 19, Score: -129.70093120676094  
Evaluating Episode: 20, Score: -134.43100161535486  
Evaluating Episode: 21, Score: -134.97019001836475  
Evaluating Episode: 22, Score: -136.57623038062243  
Evaluating Episode: 23, Score: -132.54006844867035  
Evaluating Episode: 24, Score: -126.64808646570557  
Evaluating Episode: 25, Score: -128.89157275584276  
Evaluating Episode: 26, Score: -138.70222188866842  
Evaluating Episode: 27, Score: -131.05148779165395  
Evaluating Episode: 28, Score: -137.23702031555644  
Evaluating Episode: 29, Score: -140.71535576416153  
Evaluating Episode: 30, Score: -129.22289055103434  
Evaluating Episode: 31, Score: -397.50110375294435  
Evaluating Episode: 32, Score: -6.4150585316396205  
Evaluating Episode: 33, Score: -142.93161769168566  
Evaluating Episode: 34, Score: -272.0509420418009  
Evaluating Episode: 35, Score: -131.69010376020742  
Evaluating Episode: 36, Score: -142.65470395516948  
Evaluating Episode: 37, Score: -260.55674456905916  
Evaluating Episode: 38, Score: -131.67636490909297  
Evaluating Episode: 39, Score: -366.5728410478583  
Evaluating Episode: 40, Score: -146.62864227917586  
Evaluating Episode: 41, Score: -243.03405605838668  
Evaluating Episode: 42, Score: -135.1980300079443  
Evaluating Episode: 43, Score: -127.69355206838236  
Evaluating Episode: 44, Score: -131.16216609004655  
Evaluating Episode: 45, Score: -147.12667279297864  
Evaluating Episode: 46, Score: -132.31846448384954  
Evaluating Episode: 47, Score: -268.08578306608825  
Evaluating Episode: 48, Score: -135.18250999508635  
Evaluating Episode: 49, Score: -7.894984536093582  
Evaluating Episode: 50, Score: -128.05266367014667  
Evaluating Episode: 51, Score: -145.10997135643865  
Evaluating Episode: 52, Score: -6.307511442206271  
Evaluating Episode: 53, Score: -146.70673507953794  
Evaluating Episode: 54, Score: -258.25634064733003  
Evaluating Episode: 55, Score: -392.41895958831503  
Evaluating Episode: 56, Score: -256.83043566708653  
Evaluating Episode: 57, Score: -140.62454048412212  
Evaluating Episode: 58, Score: -255.98136204837658  
Evaluating Episode: 59, Score: -271.7765475943526  
Evaluating Episode: 60, Score: -133.23808325439626  
Evaluating Episode: 61, Score: -141.12372662555478  
Evaluating Episode: 62, Score: -137.50976342084115  
Evaluating Episode: 63, Score: -132.62034273864901  
Evaluating Episode: 64, Score: -286.69299653916556  
Evaluating Episode: 65, Score: -124.30356347446474  
Evaluating Episode: 66, Score: -126.8487568159161

```
Evaluating Episode: 67, Score: -131.91656468265438
Evaluating Episode: 68, Score: -360.8689051934147
Evaluating Episode: 69, Score: -130.83897952388438
Evaluating Episode: 70, Score: -145.5641830928525
Evaluating Episode: 71, Score: -134.1992453211775
Evaluating Episode: 72, Score: -255.47410375973615
Evaluating Episode: 73, Score: -137.0809743920109
Evaluating Episode: 74, Score: -365.85899997343995
Evaluating Episode: 75, Score: -132.3147055556645
Evaluating Episode: 76, Score: -259.41459354046134
Evaluating Episode: 77, Score: -134.9776676323518
Evaluating Episode: 78, Score: -248.49733361455407
Evaluating Episode: 79, Score: -260.46674661687996
Evaluating Episode: 80, Score: -258.62317537411144
Evaluating Episode: 81, Score: -24.294087771305993
Evaluating Episode: 82, Score: -131.19095588168267
Evaluating Episode: 83, Score: -132.17037066382585
Evaluating Episode: 84, Score: -129.6990715536944
Evaluating Episode: 85, Score: -6.2986513883668165
Evaluating Episode: 86, Score: -478.2610704031582
Evaluating Episode: 87, Score: -142.3135551850051
Evaluating Episode: 88, Score: -197.81343548338955
Evaluating Episode: 89, Score: -6.363592572819046
Evaluating Episode: 90, Score: -270.8212461411048
Evaluating Episode: 91, Score: -257.9259828092839
Evaluating Episode: 92, Score: -259.80948411676826
Evaluating Episode: 93, Score: -133.55958352430292
Evaluating Episode: 94, Score: -140.50368114648504
Evaluating Episode: 95, Score: -138.2853680416493
Evaluating Episode: 96, Score: -351.7220932254341
Evaluating Episode: 97, Score: -291.3271589279943
Evaluating Episode: 98, Score: -245.17716921635073
Evaluating Episode: 99, Score: -140.97949187985697
Evaluating Episode: 100, Score: -249.5762971084918
Final Score: -185.0994703336971
```

Observations:

When comparing the training graphs of an epsilon-greedy model with gradient clipping to a simpler DQN, we can see more stabilisation.

### 1. Mitigation of Catastrophic Forgetting:

The training graph demonstrates improved retention of previously learned information, shown by the fact that there were no sudden and large decreases in performance.

### 2. Reduced Fluctuation of Scores:

While fluctuating at the beginning due to high epsilon resulting in more random observations, as the model takes over it consistently increases the performance and achieves a higher baseline due to more exploration.

## Enhancing Action Space

In a previous observation, I noted how the restrictive action space has an impact on the model's ability to balance the pendulum when it is above the fixed point. When the only options close to 0 are 0.5 and -0.5, the model, unable to apply a force like 0.1 or something small, is forced to repeatedly exert 0.5 and -0.5 on both sides to balance it. To give the model more choices especially around 0, I plan to enhance the action array using 2 strategies:

1. Increasing the size of action array: Increase number of options from 8 to 20. This could increase training time but could hit a higher baseline
2. Increase the concentration of values around 0: This is to give model more control when pendulum is almost balanced. I could use various distributions like quadratic or exponential. Here, I will go with exponential.

This only requires editing the DQNAgent

```
In [61]: class DQNAgentEpsGreedyAS:
    def __init__(self):
        self.steps = 0
        self.memory = Memory(10000)

        # Greedy Epsilon Hyperparams
        self.epsilon = 1.0
        self.epsilon_decay = 0.98
        self.epsilon_min = 0.001

        self.batch_size = 256
        self.adamlr = 0.001
        self.gamma = 0.98

        self.numOfAction = 20
        self.action_array = np.concatenate((( -1 * ( np.exp(np.linspace(0, 1.09, 10)) - 1) ) / 10) + 0.5, axis=0)

        self.policy_net = MLP(3, self.numOfAction, self.adamlr)

    def transform_choice(self, num):
        return self.action_array[int(num)]

    def action(self, state):
        sample = np.random.rand()

        if sample > self.epsilon:
            with torch.no_grad():
                choice = float(torch.argmax(self.policy_net(state)).numpy())
                action = self.transform_choice(choice)
        else:
            choice = np.random.choice([n for n in range(self.numOfAction)])
            action = self.transform_choice(choice)

        return choice, action

    def optimize(self):
        if len(self.memory) < self.batch_size:
            return

        batch = self.memory.sample(self.batch_size)
        state, action, reward, next_state = batch
        action = action.type(torch.int64)

        # target
        with torch.no_grad():    # parameters from previous iteration are held fixed
            yi = reward + self.gamma * self.policy_net(next_state).max(1)[0].unsqueeze(1)

        # Q(phi_i, a_i, theta)
        current_state_action = self.policy_net(state).gather(1,action)

        # Back_propagation
        loss = (yi - current_state_action) ** 2
        self.policy_net.optimizer.zero_grad()
        loss.mean().backward()
```

```
nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
self.policy_net.optimizer.step()
```

```
In [62]: asAgent, as_score = train_model_eps(DQNAgentEpsGreedyAS,300)
plot_tuples([(eps_score,"base"),(as_score,"action")],"Compare Increased Action Array")
```

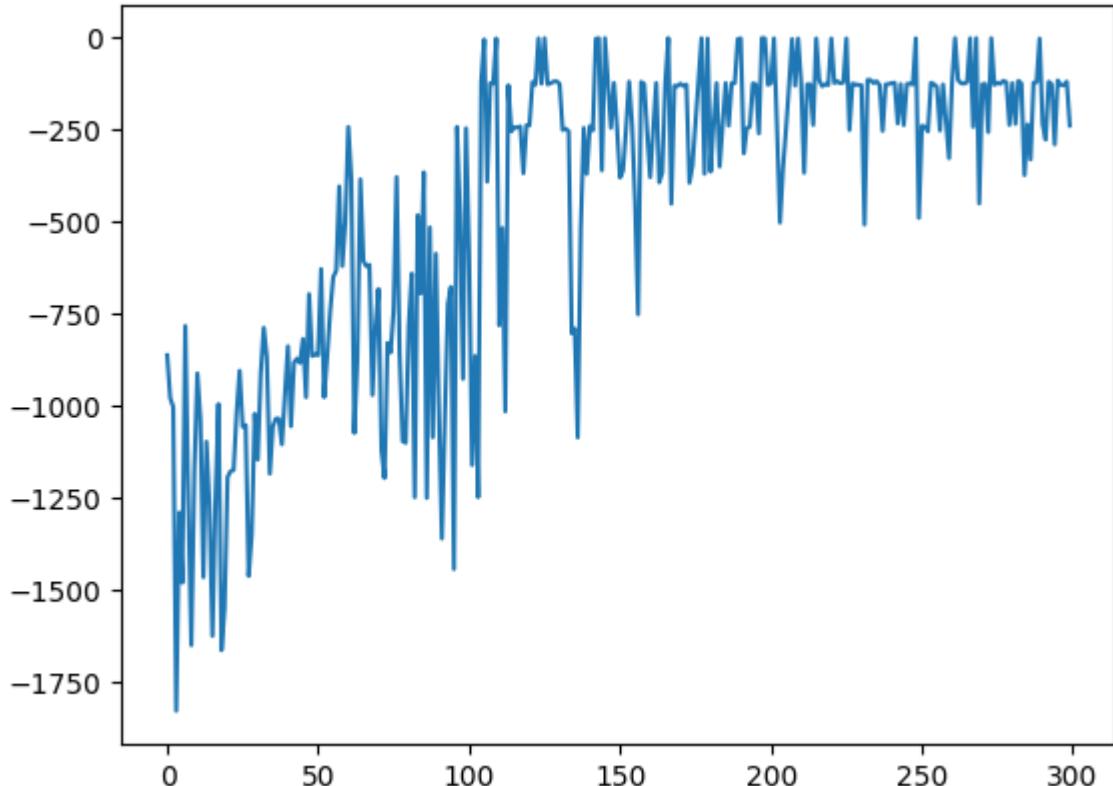
Episode: 0, Score: -861.8323603056687, Threshold: 1.0  
Episode: 1, Score: -977.9632413882897, Threshold: 0.98  
Episode: 2, Score: -1002.2041243995841, Threshold: 0.9603999999999999  
Episode: 3, Score: -1826.5246543152543, Threshold: 0.9411919999999999  
Episode: 4, Score: -1288.503634732935, Threshold: 0.9223681599999999  
Episode: 5, Score: -1478.970993307426, Threshold: 0.9039207967999998  
Episode: 6, Score: -782.3274432344673, Threshold: 0.8858423808639998  
Episode: 7, Score: -1250.884811600396, Threshold: 0.8681255332467198  
Episode: 8, Score: -1648.4508045805835, Threshold: 0.8507630225817854  
Episode: 9, Score: -1175.7128593906061, Threshold: 0.8337477621301497  
Episode: 10, Score: -911.1186058033562, Threshold: 0.8170728068875467  
Episode: 11, Score: -1034.6247763649965, Threshold: 0.8007313507497957  
Episode: 12, Score: -1464.0779385946246, Threshold: 0.7847167237347998  
Episode: 13, Score: -1096.447503878644, Threshold: 0.7690223892601038  
Episode: 14, Score: -1269.0280285675835, Threshold: 0.7536419414749017  
Episode: 15, Score: -1623.2806428086988, Threshold: 0.7385691026454037  
Episode: 16, Score: -1248.2815267479396, Threshold: 0.7237977205924956  
Episode: 17, Score: -994.2486013336264, Threshold: 0.7093217661806457  
Episode: 18, Score: -1662.1429165701006, Threshold: 0.6951353308570327  
Episode: 19, Score: -1548.8091705270713, Threshold: 0.6812326242398921  
Episode: 20, Score: -1192.7378572325988, Threshold: 0.6676079717550942  
Episode: 21, Score: -1176.1876110135274, Threshold: 0.6542558123199923  
Episode: 22, Score: -1173.7644032277567, Threshold: 0.6411706960735924  
Episode: 23, Score: -1018.2137410648238, Threshold: 0.6283472821521205  
Episode: 24, Score: -903.8758574960482, Threshold: 0.6157803365090782  
Episode: 25, Score: -1055.8923744255203, Threshold: 0.6034647297788965  
Episode: 26, Score: -1051.3890355658853, Threshold: 0.5913954351833186  
Episode: 27, Score: -1460.7604902988487, Threshold: 0.5795675264796523  
Episode: 28, Score: -1348.5887688516852, Threshold: 0.5679761759500592  
Episode: 29, Score: -1020.6145522942398, Threshold: 0.5566166524310581  
Episode: 30, Score: -1144.992197006925, Threshold: 0.5454843193824369  
Episode: 31, Score: -909.6545281974256, Threshold: 0.5345746329947881  
Episode: 32, Score: -787.3015145664804, Threshold: 0.5238831403348924  
Episode: 33, Score: -867.1966014328134, Threshold: 0.5134054775281945  
Episode: 34, Score: -1183.2137442826752, Threshold: 0.5031373679776306  
Episode: 35, Score: -1054.7374018939367, Threshold: 0.493074620618078  
Episode: 36, Score: -1033.8572832184145, Threshold: 0.48321312820571644  
Episode: 37, Score: -1033.9439425642177, Threshold: 0.4735488656416021  
Episode: 38, Score: -1103.7476332359486, Threshold: 0.46407788832877006  
Episode: 39, Score: -968.6618743318384, Threshold: 0.45479633056219465  
Episode: 40, Score: -837.5716268620685, Threshold: 0.44570040395095073  
Episode: 41, Score: -1054.652895275203, Threshold: 0.4367863958719317  
Episode: 42, Score: -882.1004226434931, Threshold: 0.42805066795449304  
Episode: 43, Score: -871.0893597527064, Threshold: 0.41948965459540316  
Episode: 44, Score: -882.7061456375953, Threshold: 0.4110998615034951  
Episode: 45, Score: -817.5061560115998, Threshold: 0.4028778642734252  
Episode: 46, Score: -975.8372775120001, Threshold: 0.39482030698795667  
Episode: 47, Score: -695.951226374432, Threshold: 0.38692390084819756  
Episode: 48, Score: -863.8485030244536, Threshold: 0.3791854228312336  
Episode: 49, Score: -860.0456749143441, Threshold: 0.37160171437460887  
Episode: 50, Score: -861.7922664827405, Threshold: 0.3641696800871167  
Episode: 51, Score: -627.9848477358674, Threshold: 0.35688628648537435  
Episode: 52, Score: -975.7994973842029, Threshold: 0.34974856075566685  
Episode: 53, Score: -864.5816672098397, Threshold: 0.3427535895405535  
Episode: 54, Score: -742.7598288305642, Threshold: 0.33589851774974244  
Episode: 55, Score: -650.2934235328439, Threshold: 0.3291805473947476  
Episode: 56, Score: -632.0296955558381, Threshold: 0.32259693644685267  
Episode: 57, Score: -403.84217899051106, Threshold: 0.3161449977179156  
Episode: 58, Score: -620.0625562514559, Threshold: 0.3098220977635573  
Episode: 59, Score: -503.03214192226835, Threshold: 0.30362565580828615  
Episode: 60, Score: -242.6105773206881, Threshold: 0.2975531426921204  
Episode: 61, Score: -386.03908792784944, Threshold: 0.291602079838278  
Episode: 62, Score: -1072.8868558876004, Threshold: 0.2857700382415124  
Episode: 63, Score: -869.7752580869628, Threshold: 0.2800546374766822  
Episode: 64, Score: -384.7634178856804, Threshold: 0.27445354472714856  
Episode: 65, Score: -606.207504524203, Threshold: 0.2689644738326056

Episode: 66, Score: -620.9170526678124, Threshold: 0.26358518435595346  
Episode: 67, Score: -616.5999441196716, Threshold: 0.25831348066883436  
Episode: 68, Score: -969.7004500967312, Threshold: 0.2531472110554577  
Episode: 69, Score: -782.40715496791, Threshold: 0.24808426683434853  
Episode: 70, Score: -682.1572163428075, Threshold: 0.24312258149766156  
Episode: 71, Score: -1120.104868484195, Threshold: 0.2382601298677083  
Episode: 72, Score: -1195.1059967457547, Threshold: 0.23349492727035415  
Episode: 73, Score: -828.3105087350995, Threshold: 0.22882502872494706  
Episode: 74, Score: -854.683011426195, Threshold: 0.22424852815044813  
Episode: 75, Score: -741.0433058713368, Threshold: 0.21976355758743915  
Episode: 76, Score: -378.565999777137, Threshold: 0.21536828643569036  
Episode: 77, Score: -859.3842178976888, Threshold: 0.21106092070697655  
Episode: 78, Score: -1094.9263599675605, Threshold: 0.20683970229283702  
Episode: 79, Score: -1099.4728194073034, Threshold: 0.20270290824698028  
Episode: 80, Score: -784.4892098515404, Threshold: 0.19864885008204067  
Episode: 81, Score: -640.4263224779465, Threshold: 0.19467587308039985  
Episode: 82, Score: -1247.1567732523797, Threshold: 0.19078235561879187  
Episode: 83, Score: -481.921513223788, Threshold: 0.18696670850641603  
Episode: 84, Score: -694.2811265510231, Threshold: 0.18322737433628772  
Episode: 85, Score: -365.8097955287181, Threshold: 0.17956282684956196  
Episode: 86, Score: -1249.2530454800656, Threshold: 0.1759715703125707  
Episode: 87, Score: -515.1722622410963, Threshold: 0.1724521389063193  
Episode: 88, Score: -1084.4444234513442, Threshold: 0.16900309612819292  
Episode: 89, Score: -586.5604492348621, Threshold: 0.16562303420562907  
Episode: 90, Score: -1012.5148185690089, Threshold: 0.16231057352151648  
Episode: 91, Score: -1359.164349084, Threshold: 0.15906436205108615  
Episode: 92, Score: -1023.7250147302448, Threshold: 0.15588307481006441  
Episode: 93, Score: -722.4388836920918, Threshold: 0.15276541331386312  
Episode: 94, Score: -676.6202115152054, Threshold: 0.14971010504758586  
Episode: 95, Score: -1442.0071689843683, Threshold: 0.14671590294663414  
Episode: 96, Score: -242.57084748022956, Threshold: 0.14378158488770146  
Episode: 97, Score: -481.22785788316713, Threshold: 0.14090595318994742  
Episode: 98, Score: -925.8324103320635, Threshold: 0.13808783412614847  
Episode: 99, Score: -246.22377719449403, Threshold: 0.13532607744362551  
Episode: 100, Score: -585.2707761985345, Threshold: 0.132619555894753  
Episode: 101, Score: -1159.6661048749975, Threshold: 0.12996716477685794  
Episode: 102, Score: -863.8637025086716, Threshold: 0.12736782148132078  
Episode: 103, Score: -1247.1081340629185, Threshold: 0.12482046505169436  
Episode: 104, Score: -121.79177646668748, Threshold: 0.12232405575066048  
Episode: 105, Score: -5.443028504330161, Threshold: 0.11987757463564727  
Episode: 106, Score: -391.2381889563042, Threshold: 0.11748002314293432  
Episode: 107, Score: -123.73583104192609, Threshold: 0.11513042268007563  
Episode: 108, Score: -125.19966014081434, Threshold: 0.11282781422647412  
Episode: 109, Score: -2.6747787694751395, Threshold: 0.11057125794194463  
Episode: 110, Score: -780.8976773031503, Threshold: 0.10835983278310574  
Episode: 111, Score: -516.6795017761061, Threshold: 0.10619263612744362  
Episode: 112, Score: -1014.3069356741621, Threshold: 0.10406878340489474  
Episode: 113, Score: -129.54587534712113, Threshold: 0.10198740773679685  
Episode: 114, Score: -254.67507993620814, Threshold: 0.09994765958206091  
Episode: 115, Score: -243.78619671723186, Threshold: 0.0979487063904197  
Episode: 116, Score: -244.152602977318, Threshold: 0.0959897322626113  
Episode: 117, Score: -241.21893932829025, Threshold: 0.09406993761735907  
Episode: 118, Score: -368.06552555815125, Threshold: 0.0921885388650119  
Episode: 119, Score: -237.8313222443731, Threshold: 0.09034476808771166  
Episode: 120, Score: -238.4163707950999, Threshold: 0.08853787272595742  
Episode: 121, Score: -120.65654448773797, Threshold: 0.08676711527143827  
Episode: 122, Score: -127.6323251156657, Threshold: 0.0850317729660095  
Episode: 123, Score: -2.255668209334827, Threshold: 0.08333113750668932  
Episode: 124, Score: -125.26422043846013, Threshold: 0.08166451475655553  
Episode: 125, Score: -1.8211852497837213, Threshold: 0.08003122446142442  
Episode: 126, Score: -127.06501129493249, Threshold: 0.07843059997219594  
Episode: 127, Score: -124.0635485985001, Threshold: 0.07686198797275202  
Episode: 128, Score: -119.67269310178384, Threshold: 0.07532474821329699  
Episode: 129, Score: -118.79714685760582, Threshold: 0.07381825324903105  
Episode: 130, Score: -125.75772808484118, Threshold: 0.07234188818405042  
Episode: 131, Score: -250.98734212492636, Threshold: 0.07089505042036941

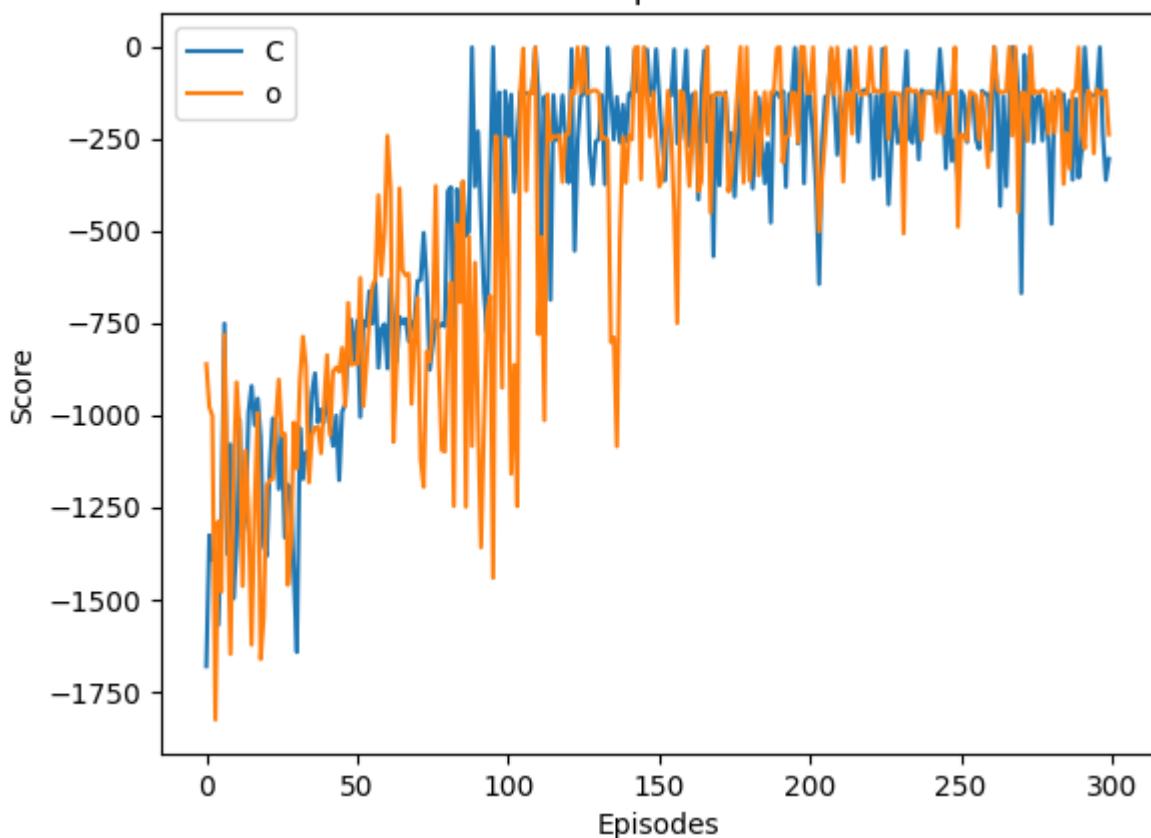
Episode: 132, Score: -247.5558074455451, Threshold: 0.06947714941196202  
Episode: 133, Score: -254.31648966122765, Threshold: 0.06808760642372277  
Episode: 134, Score: -802.6030923060537, Threshold: 0.06672585429524831  
Episode: 135, Score: -789.3366605581604, Threshold: 0.06539133720934334  
Episode: 136, Score: -1085.0014814441097, Threshold: 0.06408351046515648  
Episode: 137, Score: -519.031063662025, Threshold: 0.06280184025585335  
Episode: 138, Score: -245.22165842463724, Threshold: 0.061545803450736285  
Episode: 139, Score: -370.1201877161884, Threshold: 0.060314887381721555  
Episode: 140, Score: -240.46043960522383, Threshold: 0.05910858963408712  
Episode: 141, Score: -252.3931623474104, Threshold: 0.05792641784140538  
Episode: 142, Score: -3.372617503065919, Threshold: 0.05676788948457727  
Episode: 143, Score: -1.9130342868428611, Threshold: 0.055632531694885724  
Episode: 144, Score: -360.3808927681766, Threshold: 0.054519881060988006  
Episode: 145, Score: -1.7245199277760754, Threshold: 0.05342948343976824  
Episode: 146, Score: -121.71777249708052, Threshold: 0.052360893770972874  
Episode: 147, Score: -245.17998877096773, Threshold: 0.05131367589555342  
Episode: 148, Score: -123.26307995789846, Threshold: 0.05028740237764235  
Episode: 149, Score: -244.40075986598828, Threshold: 0.0492816543300895  
Episode: 150, Score: -380.27064018460106, Threshold: 0.04829602124348771  
Episode: 151, Score: -361.13107360970173, Threshold: 0.04733010081861796  
Episode: 152, Score: -248.1130294285424, Threshold: 0.0463834988022456  
Episode: 153, Score: -119.76939692016346, Threshold: 0.04545582882620069  
Episode: 154, Score: -236.64404033603827, Threshold: 0.04454671224967667  
Episode: 155, Score: -447.35651876650263, Threshold: 0.043655778004683135  
Episode: 156, Score: -750.9140245763463, Threshold: 0.04278266244458947  
Episode: 157, Score: -120.65578666094291, Threshold: 0.04192700919569768  
Episode: 158, Score: -125.81969829960119, Threshold: 0.04108846901178373  
Episode: 159, Score: -257.5897847228399, Threshold: 0.04026669963154805  
Episode: 160, Score: -378.9442718626723, Threshold: 0.03946136563891709  
Episode: 161, Score: -246.45501221700914, Threshold: 0.03867213832613875  
Episode: 162, Score: -123.45904402187767, Threshold: 0.037898695559615975  
Episode: 163, Score: -392.77136547379683, Threshold: 0.03714072164842366  
Episode: 164, Score: -366.1018256486856, Threshold: 0.03639790721545518  
Episode: 165, Score: -125.60691745591821, Threshold: 0.035669949071146075  
Episode: 166, Score: -1.4616791089985168, Threshold: 0.034956550089723155  
Episode: 167, Score: -451.05432320345835, Threshold: 0.03425741908792869  
Episode: 168, Score: -130.94168516906407, Threshold: 0.03357227070617012  
Episode: 169, Score: -131.621993841618, Threshold: 0.032900825292046715  
Episode: 170, Score: -125.34060637452747, Threshold: 0.03224280878620578  
Episode: 171, Score: -131.2292607331379, Threshold: 0.031597952610481664  
Episode: 172, Score: -127.73350993119202, Threshold: 0.03096599355827203  
Episode: 173, Score: -394.0159822972737, Threshold: 0.030346673687106588  
Episode: 174, Score: -347.7804034633896, Threshold: 0.029739740213364455  
Episode: 175, Score: -233.9068360384449, Threshold: 0.029144945409097164  
Episode: 176, Score: -117.35223681250989, Threshold: 0.02856204650091522  
Episode: 177, Score: -1.8210538059291301, Threshold: 0.027990805570896914  
Episode: 178, Score: -369.64016394962204, Threshold: 0.027430989459478975  
Episode: 179, Score: -2.665373408281109, Threshold: 0.026882369670289397  
Episode: 180, Score: -363.9504076635776, Threshold: 0.026344722276883607  
Episode: 181, Score: -234.19434774390467, Threshold: 0.025817827831345935  
Episode: 182, Score: -123.85407263436377, Threshold: 0.025301471274719014  
Episode: 183, Score: -349.95375271470465, Threshold: 0.024795441849224633  
Episode: 184, Score: -233.47522929382237, Threshold: 0.02429953301224014  
Episode: 185, Score: -123.6316974992735, Threshold: 0.02381354235199534  
Episode: 186, Score: -239.19087027802192, Threshold: 0.02333727150495543  
Episode: 187, Score: -123.82620713657671, Threshold: 0.02287052607485632  
Episode: 188, Score: -124.09677263090768, Threshold: 0.022413115553359196  
Episode: 189, Score: -4.553901821948576, Threshold: 0.021964853242292012  
Episode: 190, Score: -2.144403792509429, Threshold: 0.02152555617744617  
Episode: 191, Score: -314.13376342286796, Threshold: 0.021095045053897248  
Episode: 192, Score: -246.50825662869354, Threshold: 0.020673144152819303  
Episode: 193, Score: -241.81842935098052, Threshold: 0.02025968126976292  
Episode: 194, Score: -124.61429863248189, Threshold: 0.01985448764436766  
Episode: 195, Score: -126.90140210421256, Threshold: 0.019457397891480308  
Episode: 196, Score: -260.17425878249037, Threshold: 0.0190682499336507  
Episode: 197, Score: -1.3986304714990476, Threshold: 0.018686884934977686

Episode: 198, Score: -2.392638125917986, Threshold: 0.01831314723627813  
Episode: 199, Score: -128.1654996064234, Threshold: 0.017946884291552567  
Episode: 200, Score: -123.76306990590722, Threshold: 0.017587946605721515  
Episode: 201, Score: -2.2839178247173777, Threshold: 0.017236187673607083  
Episode: 202, Score: -242.36964580299787, Threshold: 0.016891463920134943  
Episode: 203, Score: -502.11082110415964, Threshold: 0.016553634641732245  
Episode: 204, Score: -354.2270817190854, Threshold: 0.0162225619488976  
Episode: 205, Score: -247.7866704846459, Threshold: 0.015898110709919647  
Episode: 206, Score: -130.51854503388762, Threshold: 0.015580148495721254  
Episode: 207, Score: -4.21753486908841, Threshold: 0.015268545525806828  
Episode: 208, Score: -130.65998420064045, Threshold: 0.014963174615290691  
Episode: 209, Score: -2.7576799006060546, Threshold: 0.014663911122984877  
Episode: 210, Score: -121.61906053228302, Threshold: 0.01437063290052518  
Episode: 211, Score: -366.8581637743323, Threshold: 0.014083220242514675  
Episode: 212, Score: -127.13532963640239, Threshold: 0.013801555837664382  
Episode: 213, Score: -126.94633625337707, Threshold: 0.013525524720911095  
Episode: 214, Score: -237.10787160032928, Threshold: 0.013255014226492874  
Episode: 215, Score: -2.5794896297969183, Threshold: 0.012989913941963017  
Episode: 216, Score: -115.79882327617335, Threshold: 0.012730115663123756  
Episode: 217, Score: -132.2839607615604, Threshold: 0.012475513349861281  
Episode: 218, Score: -126.82010064738131, Threshold: 0.012226003082864055  
Episode: 219, Score: -129.0200719410475, Threshold: 0.011981483021206774  
Episode: 220, Score: -2.655413522623235, Threshold: 0.011741853360782638  
Episode: 221, Score: -123.30001416076196, Threshold: 0.011507016293566985  
Episode: 222, Score: -119.247704278374, Threshold: 0.011276875967695645  
Episode: 223, Score: -124.16915449403903, Threshold: 0.011051338448341731  
Episode: 224, Score: -121.93202688157184, Threshold: 0.010830311679374895  
Episode: 225, Score: -2.561838091712972, Threshold: 0.010613705445787398  
Episode: 226, Score: -251.08818272565472, Threshold: 0.01040143133687165  
Episode: 227, Score: -125.1302329617253, Threshold: 0.010193402710134216  
Episode: 228, Score: -127.40428995450304, Threshold: 0.009989534655931531  
Episode: 229, Score: -128.41441635576814, Threshold: 0.0097897439628129  
Episode: 230, Score: -129.69930282213474, Threshold: 0.009593949083556642  
Episode: 231, Score: -507.32458321876777, Threshold: 0.009402070101885509  
Episode: 232, Score: -115.029791010494, Threshold: 0.009214028699847799  
Episode: 233, Score: -115.6302700787369, Threshold: 0.009029748125850843  
Episode: 234, Score: -123.52003923578637, Threshold: 0.008849153163333826  
Episode: 235, Score: -118.77940879300195, Threshold: 0.008672170100067149  
Episode: 236, Score: -128.91152786189312, Threshold: 0.008498726698065806  
Episode: 237, Score: -253.36866335797578, Threshold: 0.00832875216410449  
Episode: 238, Score: -127.36192454165067, Threshold: 0.0081621771208224  
Episode: 239, Score: -126.33683291058176, Threshold: 0.007998933578405951  
Episode: 240, Score: -123.87642340974377, Threshold: 0.007838954906837833  
Episode: 241, Score: -122.77451825851108, Threshold: 0.007682175808701076  
Episode: 242, Score: -233.23315097392853, Threshold: 0.007528532292527055  
Episode: 243, Score: -126.36715328327905, Threshold: 0.007377961646676513  
Episode: 244, Score: -237.51827532647238, Threshold: 0.007230402413742983  
Episode: 245, Score: -126.43143893688497, Threshold: 0.007085794365468123  
Episode: 246, Score: -125.44206649851293, Threshold: 0.00694407847815876  
Episode: 247, Score: -127.18741838376394, Threshold: 0.0068051969085955845  
Episode: 248, Score: -2.3855100162935927, Threshold: 0.006669092970423673  
Episode: 249, Score: -489.36484984815957, Threshold: 0.00653571111015199  
Episode: 250, Score: -239.4728324683565, Threshold: 0.006404996888794895  
Episode: 251, Score: -241.85627909010225, Threshold: 0.006276896951018997  
Episode: 252, Score: -254.95114898881914, Threshold: 0.006151359011998617  
Episode: 253, Score: -122.8979767824466, Threshold: 0.006028331831758644  
Episode: 254, Score: -126.63773997524991, Threshold: 0.005907765195123471  
Episode: 255, Score: -134.18898975321054, Threshold: 0.005789609891221002  
Episode: 256, Score: -252.26889198689815, Threshold: 0.005673817693396582  
Episode: 257, Score: -127.3111245373943, Threshold: 0.00556034133952865  
Episode: 258, Score: -229.63787616064633, Threshold: 0.005449134512738076  
Episode: 259, Score: -327.4974933443883, Threshold: 0.0053401518224833145  
Episode: 260, Score: -121.6745906346526, Threshold: 0.005233348786033648  
Episode: 261, Score: -1.5279584055912288, Threshold: 0.005128681810312975  
Episode: 262, Score: -115.62787842664447, Threshold: 0.005026108174106716  
Episode: 263, Score: -124.05341163998568, Threshold: 0.004925586010624581

Episode: 264, Score: -125.0621302391224, Threshold: 0.004827074290412089  
Episode: 265, Score: -120.50163275581413, Threshold: 0.004730532804603847  
Episode: 266, Score: -1.9405608544715116, Threshold: 0.00463592214851177  
Episode: 267, Score: -242.34999233554927, Threshold: 0.004543203705541535  
Episode: 268, Score: -1.6038931243941343, Threshold: 0.004452339631430704  
Episode: 269, Score: -450.1391522203416, Threshold: 0.00436329283880209  
Episode: 270, Score: -126.86415429175808, Threshold: 0.004276026982026048  
Episode: 271, Score: -125.66957137291115, Threshold: 0.004190506442385527  
Episode: 272, Score: -256.6356804315808, Threshold: 0.004106696313537816  
Episode: 273, Score: -1.7700831181399186, Threshold: 0.00402456238726706  
Episode: 274, Score: -128.35247030659397, Threshold: 0.003944071139521719  
Episode: 275, Score: -122.72983696469885, Threshold: 0.0038651897167312848  
Episode: 276, Score: -125.72474645671852, Threshold: 0.003787885922396659  
Episode: 277, Score: -117.63772669653133, Threshold: 0.003712128203948726  
Episode: 278, Score: -121.5785429037922, Threshold: 0.0036378856398697514  
Episode: 279, Score: -237.0723407696626, Threshold: 0.0035651279270723563  
Episode: 280, Score: -123.7778668055676, Threshold: 0.0034938253685309092  
Episode: 281, Score: -234.42363783777358, Threshold: 0.003423948861160291  
Episode: 282, Score: -117.41534890604422, Threshold: 0.0033554698839370852  
Episode: 283, Score: -125.19134046976959, Threshold: 0.0032883604862583436  
Episode: 284, Score: -373.19647611640767, Threshold: 0.0032225932765331767  
Episode: 285, Score: -235.44095596675055, Threshold: 0.003158141411002513  
Episode: 286, Score: -330.9494608166211, Threshold: 0.003094978582782463  
Episode: 287, Score: -122.7793069122554, Threshold: 0.0030330790111268136  
Episode: 288, Score: -121.04610175915788, Threshold: 0.0029724174309042773  
Episode: 289, Score: -2.7874775874173814, Threshold: 0.002912969082286192  
Episode: 290, Score: -235.1411433801053, Threshold: 0.002854709700640468  
Episode: 291, Score: -277.04536448818254, Threshold: 0.0027976155066276585  
Episode: 292, Score: -120.01304462335624, Threshold: 0.0027416631964951054  
Episode: 293, Score: -127.05686603823263, Threshold: 0.0026868299325652034  
Episode: 294, Score: -289.9318921490468, Threshold: 0.002633093333913899  
Episode: 295, Score: -116.42743710179931, Threshold: 0.002580431467235621  
Episode: 296, Score: -128.4256229388928, Threshold: 0.0025288228378909086  
Episode: 297, Score: -129.75843210733254, Threshold: 0.0024782463811330903  
Episode: 298, Score: -119.69071345156638, Threshold: 0.0024286814535104282  
Episode: 299, Score: -238.77235133720902, Threshold: 0.0023801078244402194

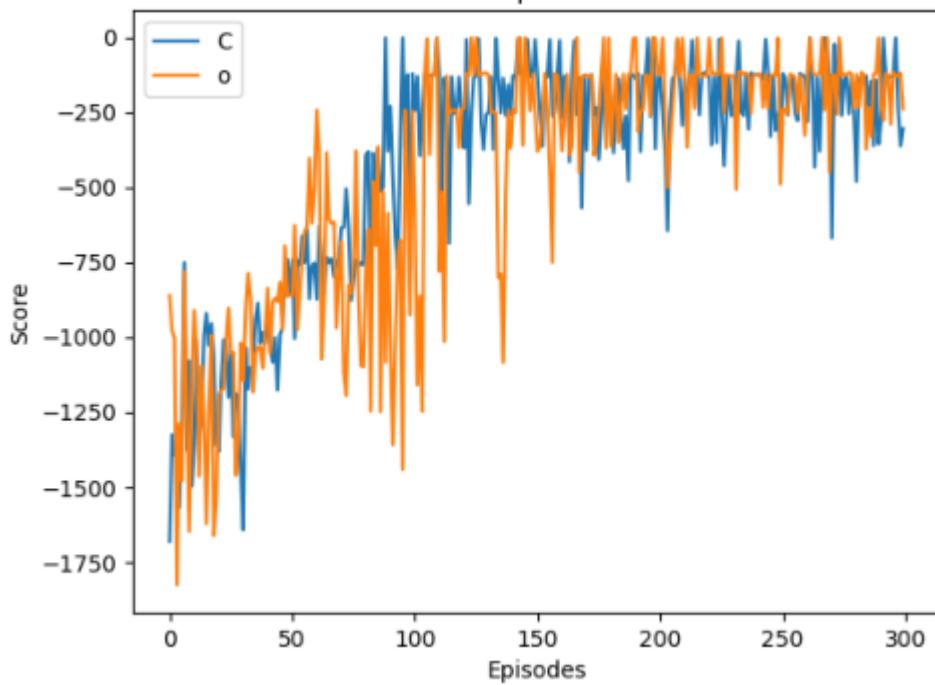


Comparison



```
In [68]: torch.save(asAgent.policy_net.state_dict(), "./rl_models/eps_action_model.pt")
showImage("./Img/RL/actionarray.png")
```

Comparison



```
In [63]: evaluate_model(asAgent, 100)
```

Evaluating Episode: 1, Score: -360.78961258968786  
Evaluating Episode: 2, Score: -122.38719155656338  
Evaluating Episode: 3, Score: -243.44224126089134  
Evaluating Episode: 4, Score: -126.52933637399833  
Evaluating Episode: 5, Score: -128.6612468782026  
Evaluating Episode: 6, Score: -3.1326407882633283  
Evaluating Episode: 7, Score: -124.75503339931856  
Evaluating Episode: 8, Score: -242.636010429231  
Evaluating Episode: 9, Score: -127.64899731619988  
Evaluating Episode: 10, Score: -125.41580794131399  
Evaluating Episode: 11, Score: -121.21512173249374  
Evaluating Episode: 12, Score: -127.78351388588779  
Evaluating Episode: 13, Score: -119.12153274595357  
Evaluating Episode: 14, Score: -121.46858733304083  
Evaluating Episode: 15, Score: -132.63969494596336  
Evaluating Episode: 16, Score: -490.04706571847294  
Evaluating Episode: 17, Score: -124.02186810245344  
Evaluating Episode: 18, Score: -254.80615549384908  
Evaluating Episode: 19, Score: -129.1419701111082  
Evaluating Episode: 20, Score: -347.2431583652955  
Evaluating Episode: 21, Score: -118.70264283759381  
Evaluating Episode: 22, Score: -326.42137964307636  
Evaluating Episode: 23, Score: -123.42088703191597  
Evaluating Episode: 24, Score: -119.69532737527608  
Evaluating Episode: 25, Score: -1.4293201252094945  
Evaluating Episode: 26, Score: -264.09992808947527  
Evaluating Episode: 27, Score: -257.09501207465274  
Evaluating Episode: 28, Score: -246.5370075405787  
Evaluating Episode: 29, Score: -126.41709005879977  
Evaluating Episode: 30, Score: -232.00862733756227  
Evaluating Episode: 31, Score: -116.54856756352862  
Evaluating Episode: 32, Score: -124.34010052612078  
Evaluating Episode: 33, Score: -1.4677410249790506  
Evaluating Episode: 34, Score: -127.14351216970752  
Evaluating Episode: 35, Score: -358.75771133218296  
Evaluating Episode: 36, Score: -125.27226885019982  
Evaluating Episode: 37, Score: -2.3295347661322587  
Evaluating Episode: 38, Score: -124.05817334163795  
Evaluating Episode: 39, Score: -125.56040004760536  
Evaluating Episode: 40, Score: -121.41750829523214  
Evaluating Episode: 41, Score: -3.6487016685491396  
Evaluating Episode: 42, Score: -367.68912579674594  
Evaluating Episode: 43, Score: -124.25336205014932  
Evaluating Episode: 44, Score: -121.39778430391299  
Evaluating Episode: 45, Score: -120.9891757140828  
Evaluating Episode: 46, Score: -118.28506043551201  
Evaluating Episode: 47, Score: -125.20817954561625  
Evaluating Episode: 48, Score: -237.16996070068848  
Evaluating Episode: 49, Score: -295.14771370890435  
Evaluating Episode: 50, Score: -242.7572218009955  
Evaluating Episode: 51, Score: -235.21521058987804  
Evaluating Episode: 52, Score: -127.40137132149403  
Evaluating Episode: 53, Score: -3.1373794009807057  
Evaluating Episode: 54, Score: -125.58187381866897  
Evaluating Episode: 55, Score: -236.7812640011918  
Evaluating Episode: 56, Score: -122.5061860652832  
Evaluating Episode: 57, Score: -240.37398518246326  
Evaluating Episode: 58, Score: -231.72064527237114  
Evaluating Episode: 59, Score: -124.14188222719245  
Evaluating Episode: 60, Score: -247.86750383967032  
Evaluating Episode: 61, Score: -121.72316209895611  
Evaluating Episode: 62, Score: -120.73106647802209  
Evaluating Episode: 63, Score: -324.22777822123595  
Evaluating Episode: 64, Score: -122.78776592224926  
Evaluating Episode: 65, Score: -120.78270645188455  
Evaluating Episode: 66, Score: -125.95698508870214

```
Evaluating Episode: 67, Score: -120.46859031742116
Evaluating Episode: 68, Score: -2.5840013495755265
Evaluating Episode: 69, Score: -126.55275130966209
Evaluating Episode: 70, Score: -125.66331517971363
Evaluating Episode: 71, Score: -117.754046063749
Evaluating Episode: 72, Score: -222.83145481096423
Evaluating Episode: 73, Score: -124.83289359591157
Evaluating Episode: 74, Score: -119.81590319997524
Evaluating Episode: 75, Score: -4.95980045265088
Evaluating Episode: 76, Score: -244.70867216113837
Evaluating Episode: 77, Score: -229.98740958034145
Evaluating Episode: 78, Score: -1.5393350351754176
Evaluating Episode: 79, Score: -125.810069810166
Evaluating Episode: 80, Score: -125.88219660584316
Evaluating Episode: 81, Score: -126.26903270405515
Evaluating Episode: 82, Score: -227.75916137718352
Evaluating Episode: 83, Score: -118.40363002800623
Evaluating Episode: 84, Score: -124.52034188931945
Evaluating Episode: 85, Score: -125.5482651942735
Evaluating Episode: 86, Score: -123.32754632398597
Evaluating Episode: 87, Score: -126.33577631138621
Evaluating Episode: 88, Score: -126.98413434126445
Evaluating Episode: 89, Score: -245.40224711309963
Evaluating Episode: 90, Score: -118.77553023772543
Evaluating Episode: 91, Score: -115.43344151808317
Evaluating Episode: 92, Score: -115.88926956168315
Evaluating Episode: 93, Score: -127.16003896980274
Evaluating Episode: 94, Score: -115.80293234616313
Evaluating Episode: 95, Score: -348.1621352477874
Evaluating Episode: 96, Score: -257.5322074135525
Evaluating Episode: 97, Score: -343.2511152518756
Evaluating Episode: 98, Score: -127.30009353606876
Evaluating Episode: 99, Score: -125.72503652589835
Evaluating Episode: 100, Score: -363.41519736325944
Final Score: -161.9948015143182
```

```
In [69]: visualize_model(asAgent,3)
```

```
Visualizing Episode: 1, Score: -126.84684030612024
Visualizing Episode: 2, Score: -1.7658373225375716
Visualizing Episode: 3, Score: -252.42271059466992
```

Observations:

- **Higher Baseline:** Based on the evaluation score, we can see tremendous improvement from -180 to -160. This is due to finer control over the inputs to environment. This was also observed in the visualisation.
- **Longer Training:** Compared to the previous model, this one takes way longer to converge to a acceptable baseline. This is because when we increase the action space, there are more things for a model to learn. Just like a human when you are given more controls to learn you would take longer.

## Enhancing Model Architecture

While the model is quite stable as of now, we seek to further stabilise it by make some architectural changes

### Double DQN

#### Problem with Simple DQN

In a standard Deep Q-Network (DQN), during the Q-value estimation for the next state, the same network is used to both select and evaluate the best action. This can lead to overestimation of Q-values, especially in scenarios where the Q-values are noisy or subject to fluctuations.

### Solution: Decouple Action Selection and Evaluation

Double DQN introduces a simple yet effective modification to the Q-value estimation process. Instead of relying on a single network for both action selection and evaluation, it maintains two separate networks: one for selecting the best action and another for evaluating that action's Q-value.

### Modifying DQN Architecture for Double DQN

Target Evaluation Network ( $Q_{\text{eval}}$ ):

- Separate neural network for evaluating Q-values of the selected actions.
- Both networks should be initialized with the same parameters.
- Parameters are periodically synchronized from the action selection network.

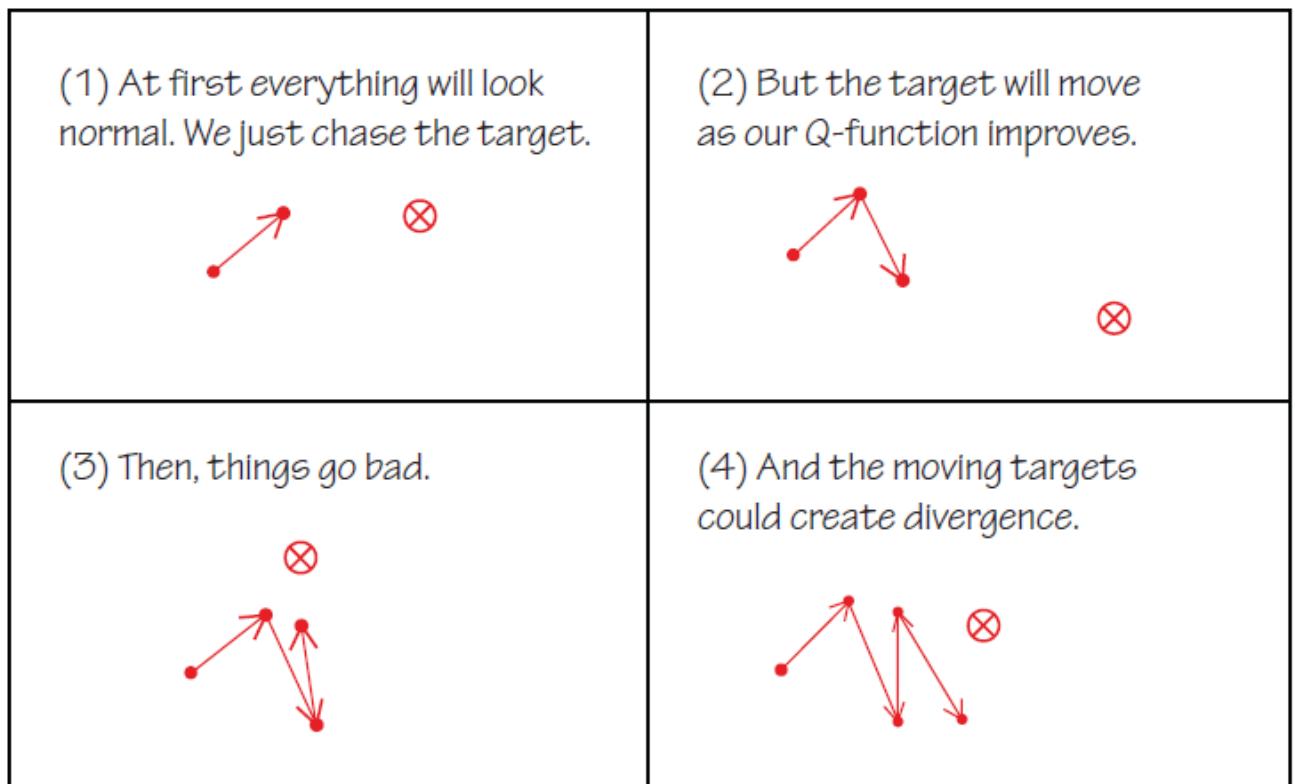
Bellman Update:

- Update the Q-value of the current state-action pair based on the reward and the discounted maximum Q-value from the target evaluation network.

### Rationale Behind using Two Networks:

Having two networks to make a decision might seem counter intuitive but it can actually stabilise training time.

Without Target Network



With Target network

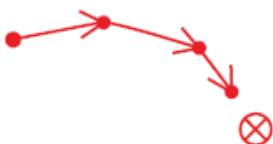
(1) Suppose we freeze the target for a few steps.



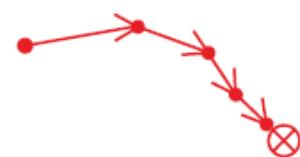
(2) That way the optimizer can make stable progress towards it.



(3) We eventually update the target, and repeat.



(4) This allows the algorithm to make stable progress.



## Loss Function

For the loss function, we will be using Huber Loss. Huber loss is calculated using the outputs from the target and policy network. Based on the Gamma parameter used, it can behave more like MSE for small errors and MAE for large errors. We will use this loss to backpropagate the policy network. We will update the target network using a weighted sum of the new policy values and the current target values

In [92]:

```
class DoubleDQNAgent:  
    def __init__(self):  
        self.steps = 0  
        self.memory = Memory(10000)  
  
        # Greedy Epsilon Hyperparams  
        self.epsilon = 1.0  
        self.epsilon_decay = 0.98  
        self.epsilon_min = 0.001  
  
        self.batch_size = 256  
        self.adamlr = 0.001  
        self.targetlr = 0.01  
        self.gamma = 0.98  
  
        self.numOfAction = 20  
        self.action_array = np.concatenate((-1 * (np.exp(np.linspace(0, 1.09, 10)) -  
  
        self.policy_net = MLP(3, self.numOfAction, self.adamlr)  
        self.target_net = MLP(3, self.numOfAction, self.adamlr)  
        self.target_net.load_state_dict(self.policy_net.state_dict())  
  
    def transform_choice(self, num):  
        return self.action_array[int(num)]  
  
    def action(self, state):  
  
        sample = np.random.rand()  
  
        if sample > self.epsilon:  
            with torch.no_grad():
```

```

        choice = float(torch.argmax(self.policy_net(state)).numpy())
        action = self.transform_choice(choice)

    else:
        choice = np.random.choice([n for n in range(self.numOfAction)])
        action = self.transform_choice(choice)

    return choice, action

def optimize(self):
    if len(self.memory) < self.batch_size:
        return

    batch = self.memory.sample(self.batch_size)
    state, action, reward, next_state = batch
    action = action.type(torch.int64)

    with torch.no_grad():
        expected_q = self.target_net(next_state).max(1)[0].unsqueeze(1)
        expected = reward + self.gamma * expected_q

    # Actual values - Policy Net
    actual = self.policy_net(state).gather(1,action)

    # Back_propagation
    loss = F.smooth_l1_loss(actual, expected)
    self.policy_net.optimizer.zero_grad()
    loss.mean().backward()
    nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
    self.policy_net.optimizer.step()

    for target_param,param in zip(self.target_net.parameters(),self.policy_net.pa
        target_param.data.copy_(target_param.data * (1.0 - self.targetlr) + param

```

In [93]: DDAgent, ddqn\_score = train\_model\_eps(DoubleDQNAgent, 300)  
plot\_tuples([(as\_score, "base"), (ddqn\_score, "action")], "Compare Double DQN")

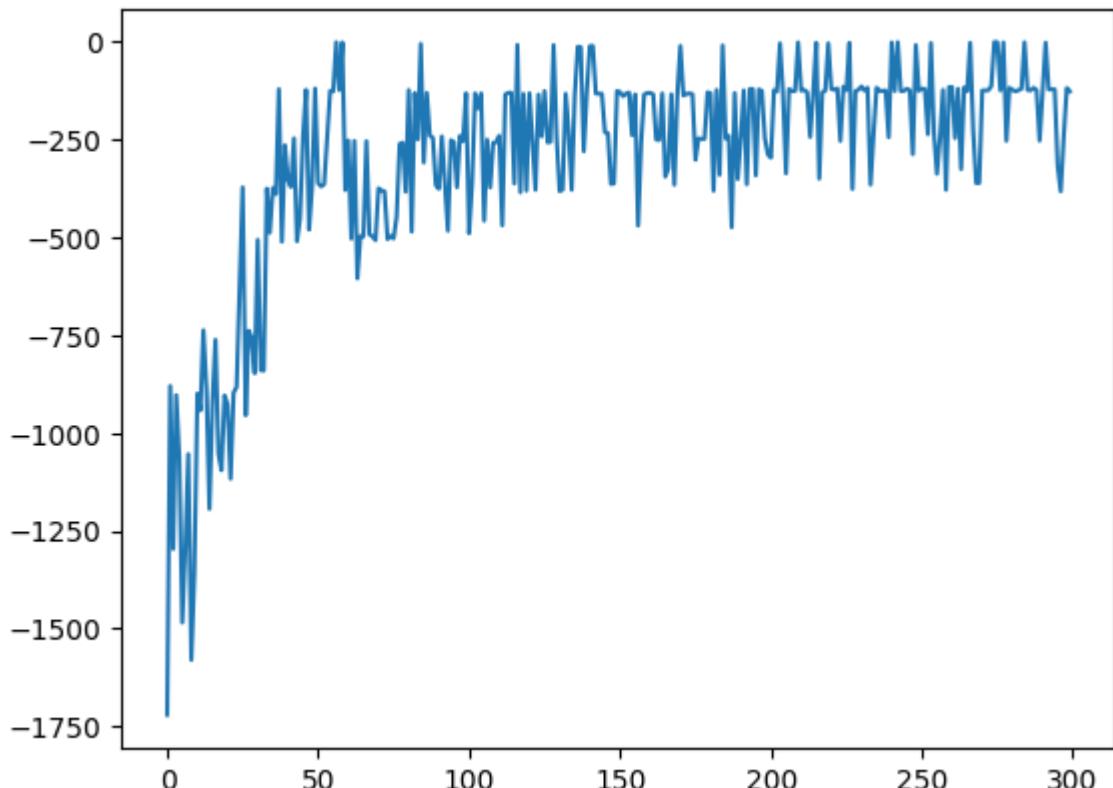
Episode: 0, Score: -1720.2270408142153, Threshold: 1.0  
Episode: 1, Score: -879.9630781382018, Threshold: 0.98  
Episode: 2, Score: -1296.3623037559014, Threshold: 0.9603999999999999  
Episode: 3, Score: -903.1763556133876, Threshold: 0.9411919999999999  
Episode: 4, Score: -1059.955271097056, Threshold: 0.9223681599999999  
Episode: 5, Score: -1484.446923097765, Threshold: 0.9039207967999998  
Episode: 6, Score: -1292.6872219982292, Threshold: 0.8858423808639998  
Episode: 7, Score: -1054.1123517864637, Threshold: 0.8681255332467198  
Episode: 8, Score: -1580.0339088996536, Threshold: 0.8507630225817854  
Episode: 9, Score: -1356.7580300869038, Threshold: 0.8337477621301497  
Episode: 10, Score: -899.3799357722331, Threshold: 0.8170728068875467  
Episode: 11, Score: -942.3626632774153, Threshold: 0.8007313507497957  
Episode: 12, Score: -737.3833816728004, Threshold: 0.7847167237347998  
Episode: 13, Score: -884.6281329942175, Threshold: 0.7690223892601038  
Episode: 14, Score: -1194.047420887284, Threshold: 0.7536419414749017  
Episode: 15, Score: -920.4576088156955, Threshold: 0.7385691026454037  
Episode: 16, Score: -761.6859141809356, Threshold: 0.7237977205924956  
Episode: 17, Score: -1053.7234719576154, Threshold: 0.7093217661806457  
Episode: 18, Score: -1095.118824614367, Threshold: 0.6951353308570327  
Episode: 19, Score: -903.9362501236927, Threshold: 0.6812326242398921  
Episode: 20, Score: -926.8803626126414, Threshold: 0.6676079717550942  
Episode: 21, Score: -1116.9881778258148, Threshold: 0.6542558123199923  
Episode: 22, Score: -897.1889886129989, Threshold: 0.6411706960735924  
Episode: 23, Score: -882.2707128578173, Threshold: 0.6283472821521205  
Episode: 24, Score: -627.993043210483, Threshold: 0.6157803365090782  
Episode: 25, Score: -372.32993821900357, Threshold: 0.6034647297788965  
Episode: 26, Score: -954.8645210504477, Threshold: 0.5913954351833186  
Episode: 27, Score: -738.7129928745436, Threshold: 0.5795675264796523  
Episode: 28, Score: -756.510902322908, Threshold: 0.5679761759500592  
Episode: 29, Score: -847.4515996764688, Threshold: 0.5566166524310581  
Episode: 30, Score: -506.35034221629843, Threshold: 0.5454843193824369  
Episode: 31, Score: -839.5060256452258, Threshold: 0.5345746329947881  
Episode: 32, Score: -840.6147836700715, Threshold: 0.5238831403348924  
Episode: 33, Score: -375.7488107600246, Threshold: 0.5134054775281945  
Episode: 34, Score: -488.01096735971623, Threshold: 0.5031373679776306  
Episode: 35, Score: -374.1858810443215, Threshold: 0.493074620618078  
Episode: 36, Score: -389.6005071457724, Threshold: 0.48321312820571644  
Episode: 37, Score: -121.43082504590069, Threshold: 0.4735488656416021  
Episode: 38, Score: -511.3070828136455, Threshold: 0.46407788832877006  
Episode: 39, Score: -264.701756795172, Threshold: 0.45479633056219465  
Episode: 40, Score: -353.30407393585295, Threshold: 0.44570040395095073  
Episode: 41, Score: -371.4789918064539, Threshold: 0.4367863958719317  
Episode: 42, Score: -247.0176833138008, Threshold: 0.42805066795449304  
Episode: 43, Score: -509.98836805922593, Threshold: 0.41948965459540316  
Episode: 44, Score: -451.4814573513522, Threshold: 0.4110998615034951  
Episode: 45, Score: -244.45620249191006, Threshold: 0.4028778642734252  
Episode: 46, Score: -123.29627580315055, Threshold: 0.39482030698795667  
Episode: 47, Score: -480.1336686006406, Threshold: 0.38692390084819756  
Episode: 48, Score: -369.57996617001584, Threshold: 0.3791854228312336  
Episode: 49, Score: -119.80276540173301, Threshold: 0.37160171437460887  
Episode: 50, Score: -361.56112471572624, Threshold: 0.3641696800871167  
Episode: 51, Score: -369.90783178793976, Threshold: 0.35688628648537435  
Episode: 52, Score: -364.4197456317267, Threshold: 0.34974856075566685  
Episode: 53, Score: -244.47821175045854, Threshold: 0.3427535895405535  
Episode: 54, Score: -125.59555008591242, Threshold: 0.33589851774974244  
Episode: 55, Score: -127.2861579158153, Threshold: 0.3291805473947476  
Episode: 56, Score: -1.4298725268660895, Threshold: 0.32259693644685267  
Episode: 57, Score: -123.88150141752965, Threshold: 0.3161449977179156  
Episode: 58, Score: -2.063744926322474, Threshold: 0.3098220977635573  
Episode: 59, Score: -379.28314111509223, Threshold: 0.30362565580828615  
Episode: 60, Score: -252.588533662851, Threshold: 0.2975531426921204  
Episode: 61, Score: -503.5797604175883, Threshold: 0.291602079838278  
Episode: 62, Score: -253.58138609925015, Threshold: 0.2857700382415124  
Episode: 63, Score: -604.894457991769, Threshold: 0.2800546374766822  
Episode: 64, Score: -496.89338578448053, Threshold: 0.27445354472714856  
Episode: 65, Score: -498.04244236299996, Threshold: 0.2689644738326056

Episode: 66, Score: -254.36664452676538, Threshold: 0.26358518435595346  
Episode: 67, Score: -492.3683418375806, Threshold: 0.25831348066883436  
Episode: 68, Score: -497.7932626499584, Threshold: 0.2531472110554577  
Episode: 69, Score: -507.3455578605054, Threshold: 0.24808426683434853  
Episode: 70, Score: -374.90581549909456, Threshold: 0.24312258149766156  
Episode: 71, Score: -380.36298366547993, Threshold: 0.2382601298677083  
Episode: 72, Score: -382.79822587237413, Threshold: 0.23349492727035415  
Episode: 73, Score: -504.716672482101, Threshold: 0.22882502872494706  
Episode: 74, Score: -496.59384001639285, Threshold: 0.22424852815044813  
Episode: 75, Score: -502.8706872343175, Threshold: 0.21976355758743915  
Episode: 76, Score: -447.72788508807514, Threshold: 0.21536828643569036  
Episode: 77, Score: -261.36832012830246, Threshold: 0.21106092070697655  
Episode: 78, Score: -258.41910241077636, Threshold: 0.20683970229283702  
Episode: 79, Score: -383.5225555001409, Threshold: 0.20270290824698028  
Episode: 80, Score: -123.62436536739094, Threshold: 0.19864885008204067  
Episode: 81, Score: -485.57915513805904, Threshold: 0.19467587308039985  
Episode: 82, Score: -131.31972463907476, Threshold: 0.19078235561879187  
Episode: 83, Score: -249.56770698838412, Threshold: 0.18696670850641603  
Episode: 84, Score: -5.835517912421252, Threshold: 0.18322737433628772  
Episode: 85, Score: -308.7687879048158, Threshold: 0.17956282684956196  
Episode: 86, Score: -130.99051406810085, Threshold: 0.1759715703125707  
Episode: 87, Score: -240.23427653234174, Threshold: 0.1724521389063193  
Episode: 88, Score: -245.4614034747305, Threshold: 0.16900309612819292  
Episode: 89, Score: -366.2784153468992, Threshold: 0.16562303420562907  
Episode: 90, Score: -375.8654971724552, Threshold: 0.16231057352151648  
Episode: 91, Score: -242.1388020023815, Threshold: 0.15906436205108615  
Episode: 92, Score: -360.3813543033183, Threshold: 0.15588307481006441  
Episode: 93, Score: -483.0231776149321, Threshold: 0.15276541331386312  
Episode: 94, Score: -252.0242477459421, Threshold: 0.14971010504758586  
Episode: 95, Score: -256.7026204333964, Threshold: 0.14671590294663414  
Episode: 96, Score: -372.037891422542, Threshold: 0.14378158488770146  
Episode: 97, Score: -240.64524970657558, Threshold: 0.14090595318994742  
Episode: 98, Score: -254.2218275416626, Threshold: 0.13808783412614847  
Episode: 99, Score: -131.9407965076822, Threshold: 0.13532607744362551  
Episode: 100, Score: -489.78047562560994, Threshold: 0.132619555894753  
Episode: 101, Score: -362.15506527863755, Threshold: 0.12996716477685794  
Episode: 102, Score: -132.3266048754926, Threshold: 0.12736782148132078  
Episode: 103, Score: -170.8496962585179, Threshold: 0.12482046505169436  
Episode: 104, Score: -132.37936095680283, Threshold: 0.12232405575066048  
Episode: 105, Score: -457.37266960010027, Threshold: 0.11987757463564727  
Episode: 106, Score: -250.59226823236753, Threshold: 0.11748002314293432  
Episode: 107, Score: -372.68697884336234, Threshold: 0.11513042268007563  
Episode: 108, Score: -258.60264647287755, Threshold: 0.11282781422647412  
Episode: 109, Score: -257.05185108769876, Threshold: 0.11057125794194463  
Episode: 110, Score: -240.44291857414038, Threshold: 0.10835983278310574  
Episode: 111, Score: -469.40539303285817, Threshold: 0.10619263612744362  
Episode: 112, Score: -135.0127861671377, Threshold: 0.10406878340489474  
Episode: 113, Score: -130.98434295551803, Threshold: 0.10198740773679685  
Episode: 114, Score: -130.54163796932136, Threshold: 0.09994765958206091  
Episode: 115, Score: -362.4966030383106, Threshold: 0.0979487063904197  
Episode: 116, Score: -8.356375512783071, Threshold: 0.0959897322626113  
Episode: 117, Score: -384.88691482161687, Threshold: 0.09406993761735907  
Episode: 118, Score: -133.32466277629769, Threshold: 0.0921885388650119  
Episode: 119, Score: -380.71998399346046, Threshold: 0.09034476808771166  
Episode: 120, Score: -131.80061223598923, Threshold: 0.08853787272595742  
Episode: 121, Score: -258.50751989528646, Threshold: 0.08676711527143827  
Episode: 122, Score: -378.66157859389955, Threshold: 0.0850317729660095  
Episode: 123, Score: -135.16183682262042, Threshold: 0.08333113750668932  
Episode: 124, Score: -242.07618129222814, Threshold: 0.08166451475655553  
Episode: 125, Score: -126.33004647557534, Threshold: 0.08003122446142442  
Episode: 126, Score: -257.36478984995404, Threshold: 0.07843059997219594  
Episode: 127, Score: -255.96456871657506, Threshold: 0.07686198797275202  
Episode: 128, Score: -8.786272554432115, Threshold: 0.07532474821329699  
Episode: 129, Score: -244.16161979808442, Threshold: 0.07381825324903105  
Episode: 130, Score: -381.8746067561275, Threshold: 0.07234188818405042  
Episode: 131, Score: -378.3818437272248, Threshold: 0.07089505042036941

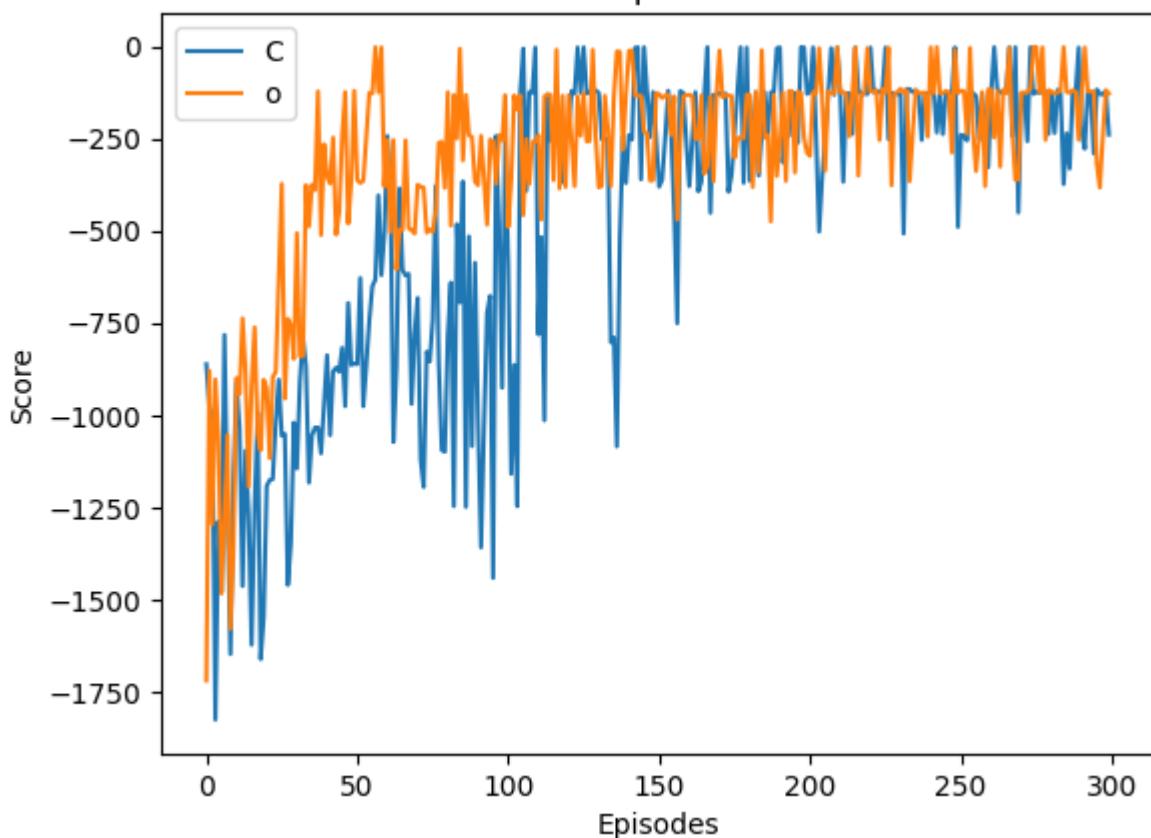
Episode: 132, Score: -133.2308364788809, Threshold: 0.06947714941196202  
Episode: 133, Score: -228.0592842139129, Threshold: 0.06808760642372277  
Episode: 134, Score: -378.662667039089, Threshold: 0.06672585429524831  
Episode: 135, Score: -137.6522086222204, Threshold: 0.06539133720934334  
Episode: 136, Score: -13.039827131294071, Threshold: 0.06408351046515648  
Episode: 137, Score: -14.15410874369942, Threshold: 0.06280184025585335  
Episode: 138, Score: -280.0861988827951, Threshold: 0.061545803450736285  
Episode: 139, Score: -130.02026966745723, Threshold: 0.060314887381721555  
Episode: 140, Score: -11.446024495577257, Threshold: 0.05910858963408712  
Episode: 141, Score: -10.982627898812524, Threshold: 0.05792641784140538  
Episode: 142, Score: -132.54999606211797, Threshold: 0.05676788948457727  
Episode: 143, Score: -132.96528224873333, Threshold: 0.055632531694885724  
Episode: 144, Score: -133.4197232995538, Threshold: 0.054519881060988006  
Episode: 145, Score: -231.48112214935819, Threshold: 0.05342948343976824  
Episode: 146, Score: -234.0779201536035, Threshold: 0.052360893770972874  
Episode: 147, Score: -363.61496312580954, Threshold: 0.05131367589555342  
Episode: 148, Score: -362.47372248597117, Threshold: 0.05028740237764235  
Episode: 149, Score: -126.23437382887795, Threshold: 0.0492816543300895  
Episode: 150, Score: -128.90689008499618, Threshold: 0.04829602124348771  
Episode: 151, Score: -139.2841559447809, Threshold: 0.04733010081861796  
Episode: 152, Score: -132.96258507926657, Threshold: 0.0463834988022456  
Episode: 153, Score: -130.17635193994818, Threshold: 0.04545582882620069  
Episode: 154, Score: -239.07694190038515, Threshold: 0.04454671224967667  
Episode: 155, Score: -135.0049431183057, Threshold: 0.043655778004683135  
Episode: 156, Score: -470.0876905593998, Threshold: 0.04278266244458947  
Episode: 157, Score: -259.5319605886851, Threshold: 0.04192700919569768  
Episode: 158, Score: -135.05141340467438, Threshold: 0.04108846901178373  
Episode: 159, Score: -131.5889190665465, Threshold: 0.04026669963154805  
Episode: 160, Score: -131.00795466799525, Threshold: 0.03946136563891709  
Episode: 161, Score: -134.79579628124984, Threshold: 0.03867213832613875  
Episode: 162, Score: -250.1180867748964, Threshold: 0.037898695559615975  
Episode: 163, Score: -251.39740716921247, Threshold: 0.03714072164842366  
Episode: 164, Score: -132.49098601143197, Threshold: 0.03639790721545518  
Episode: 165, Score: -344.9420120288003, Threshold: 0.035669949071146075  
Episode: 166, Score: -324.6119318897024, Threshold: 0.034956550089723155  
Episode: 167, Score: -134.48579307028444, Threshold: 0.03425741908792869  
Episode: 168, Score: -365.6366914509567, Threshold: 0.03357227070617012  
Episode: 169, Score: -132.85269074596818, Threshold: 0.032900825292046715  
Episode: 170, Score: -10.848198032057573, Threshold: 0.03224280878620578  
Episode: 171, Score: -137.48740074334782, Threshold: 0.031597952610481664  
Episode: 172, Score: -133.64299668866084, Threshold: 0.03096599355827203  
Episode: 173, Score: -132.40637518983226, Threshold: 0.030346673687106588  
Episode: 174, Score: -134.20824514118974, Threshold: 0.029739740213364455  
Episode: 175, Score: -301.65775686969624, Threshold: 0.029144945409097164  
Episode: 176, Score: -247.2548309861142, Threshold: 0.02856204650091522  
Episode: 177, Score: -248.5344455758051, Threshold: 0.027990805570896914  
Episode: 178, Score: -248.45536707474696, Threshold: 0.027430989459478975  
Episode: 179, Score: -129.4168324821893, Threshold: 0.026882369670289397  
Episode: 180, Score: -131.9715437556428, Threshold: 0.026344722276883607  
Episode: 181, Score: -380.9859165212889, Threshold: 0.025817827831345935  
Episode: 182, Score: -123.3644273996959, Threshold: 0.025301471274719014  
Episode: 183, Score: -340.25459226639157, Threshold: 0.024795441849224633  
Episode: 184, Score: -9.720430892670494, Threshold: 0.02429953301224014  
Episode: 185, Score: -247.38158427967178, Threshold: 0.02381354235199534  
Episode: 186, Score: -240.17903498823568, Threshold: 0.02333727150495543  
Episode: 187, Score: -474.36353887128655, Threshold: 0.02287052607485632  
Episode: 188, Score: -131.4383276806618, Threshold: 0.022413115553359196  
Episode: 189, Score: -351.1105447117424, Threshold: 0.021964853242292012  
Episode: 190, Score: -248.80331040080253, Threshold: 0.02152555617744617  
Episode: 191, Score: -124.51339696447165, Threshold: 0.021095045053897248  
Episode: 192, Score: -364.5170012198311, Threshold: 0.020673144152819303  
Episode: 193, Score: -120.98476507206533, Threshold: 0.02025968126976292  
Episode: 194, Score: -121.20968856217245, Threshold: 0.01985448764436766  
Episode: 195, Score: -341.7033400294651, Threshold: 0.019457397891480308  
Episode: 196, Score: -122.27573303245174, Threshold: 0.0190682499336507  
Episode: 197, Score: -125.98461714730209, Threshold: 0.018686884934977686

Episode: 198, Score: -247.8020973428192, Threshold: 0.01831314723627813  
Episode: 199, Score: -288.1774124394573, Threshold: 0.017946884291552567  
Episode: 200, Score: -296.60245541055525, Threshold: 0.017587946605721515  
Episode: 201, Score: -124.2828363041791, Threshold: 0.017236187673607083  
Episode: 202, Score: -127.39073074126907, Threshold: 0.016891463920134943  
Episode: 203, Score: -4.390011188466517, Threshold: 0.016553634641732245  
Episode: 204, Score: -132.6721924886568, Threshold: 0.0162225619488976  
Episode: 205, Score: -336.5892123680523, Threshold: 0.015898110709919647  
Episode: 206, Score: -121.44340832797462, Threshold: 0.015580148495721254  
Episode: 207, Score: -126.43671551073183, Threshold: 0.015268545525806828  
Episode: 208, Score: -125.97299565626689, Threshold: 0.014963174615290691  
Episode: 209, Score: -1.8422974821311717, Threshold: 0.014663911122984877  
Episode: 210, Score: -124.90973333097484, Threshold: 0.01437063290052518  
Episode: 211, Score: -123.36816990992517, Threshold: 0.014083220242514675  
Episode: 212, Score: -129.03427103226787, Threshold: 0.013801555837664382  
Episode: 213, Score: -243.4522050907464, Threshold: 0.013525524720911095  
Episode: 214, Score: -130.65712907001014, Threshold: 0.013255014226492874  
Episode: 215, Score: -2.9208975022596335, Threshold: 0.012989913941963017  
Episode: 216, Score: -350.07627908492566, Threshold: 0.012730115663123756  
Episode: 217, Score: -129.56206906551859, Threshold: 0.012475513349861281  
Episode: 218, Score: -125.38232082421305, Threshold: 0.012226003082864055  
Episode: 219, Score: -4.196553860596719, Threshold: 0.011981483021206774  
Episode: 220, Score: -120.54013359251174, Threshold: 0.011741853360782638  
Episode: 221, Score: -123.00103604050915, Threshold: 0.011507016293566985  
Episode: 222, Score: -119.63448892280921, Threshold: 0.011276875967695645  
Episode: 223, Score: -253.60714468276151, Threshold: 0.011051338448341731  
Episode: 224, Score: -114.92139647600156, Threshold: 0.010830311679374895  
Episode: 225, Score: -123.81125257087311, Threshold: 0.010613705445787398  
Episode: 226, Score: -3.057614805562858, Threshold: 0.01040143133687165  
Episode: 227, Score: -376.412183063878, Threshold: 0.010193402710134216  
Episode: 228, Score: -126.72817268445269, Threshold: 0.009989534655931531  
Episode: 229, Score: -123.00081334239948, Threshold: 0.0097897439628129  
Episode: 230, Score: -114.39052754743142, Threshold: 0.009593949083556642  
Episode: 231, Score: -124.91731493442036, Threshold: 0.009402070101885509  
Episode: 232, Score: -117.77301138762954, Threshold: 0.009214028699847799  
Episode: 233, Score: -365.1862891854223, Threshold: 0.009029748125850843  
Episode: 234, Score: -235.98807209907207, Threshold: 0.008849153163333826  
Episode: 235, Score: -117.898243524095, Threshold: 0.008672170100067149  
Episode: 236, Score: -125.50970409834105, Threshold: 0.008498726698065806  
Episode: 237, Score: -126.08095309311916, Threshold: 0.00832875216410449  
Episode: 238, Score: -124.7552949911236, Threshold: 0.0081621771208224  
Episode: 239, Score: -244.27293692132855, Threshold: 0.007998933578405951  
Episode: 240, Score: -2.016263805301267, Threshold: 0.007838954906837833  
Episode: 241, Score: -126.18240346393542, Threshold: 0.007682175808701076  
Episode: 242, Score: -1.2692417821560742, Threshold: 0.007528532292527055  
Episode: 243, Score: -126.20606882806347, Threshold: 0.007377961646676513  
Episode: 244, Score: -126.070041964022, Threshold: 0.007230402413742983  
Episode: 245, Score: -120.54254989845805, Threshold: 0.007085794365468123  
Episode: 246, Score: -124.08613653834023, Threshold: 0.00694407847815876  
Episode: 247, Score: -287.60722667375603, Threshold: 0.0068051969085955845  
Episode: 248, Score: -8.856054731884521, Threshold: 0.006669092970423673  
Episode: 249, Score: -125.60486296185006, Threshold: 0.00653571111015199  
Episode: 250, Score: -120.13524925895463, Threshold: 0.006404996888794895  
Episode: 251, Score: -121.30531420445843, Threshold: 0.006276896951018997  
Episode: 252, Score: -236.50434310408272, Threshold: 0.006151359011998617  
Episode: 253, Score: -3.6571096902241416, Threshold: 0.006028331831758644  
Episode: 254, Score: -232.8681758591961, Threshold: 0.005907765195123471  
Episode: 255, Score: -337.6810605909947, Threshold: 0.005789609891221002  
Episode: 256, Score: -245.41031496429366, Threshold: 0.005673817693396582  
Episode: 257, Score: -123.50120334885584, Threshold: 0.00556034133952865  
Episode: 258, Score: -378.82501713536357, Threshold: 0.005449134512738076  
Episode: 259, Score: -116.20778648597704, Threshold: 0.0053401518224833145  
Episode: 260, Score: -116.12756157666753, Threshold: 0.005233348786033648  
Episode: 261, Score: -246.59849933873588, Threshold: 0.005128681810312975  
Episode: 262, Score: -120.4962556906965, Threshold: 0.005026108174106716  
Episode: 263, Score: -325.7584102770444, Threshold: 0.004925586010624581

Episode: 264, Score: -117.29549371240172, Threshold: 0.004827074290412089  
Episode: 265, Score: -126.72337972529508, Threshold: 0.004730532804603847  
Episode: 266, Score: -2.8992644080995174, Threshold: 0.00463592214851177  
Episode: 267, Score: -226.8615831050894, Threshold: 0.004543203705541535  
Episode: 268, Score: -361.4160208252254, Threshold: 0.004452339631430704  
Episode: 269, Score: -361.2262638082285, Threshold: 0.00436329283880209  
Episode: 270, Score: -124.9617458917908, Threshold: 0.004276026982026048  
Episode: 271, Score: -124.44388485302912, Threshold: 0.004190506442385527  
Episode: 272, Score: -124.2234010963131, Threshold: 0.004106696313537816  
Episode: 273, Score: -113.72377358904721, Threshold: 0.00402456238726706  
Episode: 274, Score: -0.9425372719341096, Threshold: 0.003944071139521719  
Episode: 275, Score: -1.6094399602988156, Threshold: 0.0038651897167312848  
Episode: 276, Score: -124.68252712168369, Threshold: 0.003787885922396659  
Episode: 277, Score: -2.166732732217337, Threshold: 0.003712128203948726  
Episode: 278, Score: -253.26177996091764, Threshold: 0.0036378856398697514  
Episode: 279, Score: -118.10016133950748, Threshold: 0.0035651279270723563  
Episode: 280, Score: -122.9987326905198, Threshold: 0.0034938253685309092  
Episode: 281, Score: -127.06066791841121, Threshold: 0.003423948861160291  
Episode: 282, Score: -124.10083846247758, Threshold: 0.0033554698839370852  
Episode: 283, Score: -120.4308806104898, Threshold: 0.0032883604862583436  
Episode: 284, Score: -1.5479248159827355, Threshold: 0.0032225932765331767  
Episode: 285, Score: -123.25859923893012, Threshold: 0.003158141411002513  
Episode: 286, Score: -124.62907394800955, Threshold: 0.003094978582782463  
Episode: 287, Score: -118.07556804528127, Threshold: 0.0030330790111268136  
Episode: 288, Score: -123.61855833332105, Threshold: 0.0029724174309042773  
Episode: 289, Score: -252.91257351117127, Threshold: 0.002912969082286192  
Episode: 290, Score: -127.05450157454152, Threshold: 0.002854709700640468  
Episode: 291, Score: -2.5759639806418946, Threshold: 0.0027976155066276585  
Episode: 292, Score: -121.9332004294131, Threshold: 0.0027416631964951054  
Episode: 293, Score: -122.10436201997602, Threshold: 0.0026868299325652034  
Episode: 294, Score: -121.99828003501452, Threshold: 0.002633093333913899  
Episode: 295, Score: -317.9722323650689, Threshold: 0.002580431467235621  
Episode: 296, Score: -382.4309976547135, Threshold: 0.0025288228378909086  
Episode: 297, Score: -241.89491398962667, Threshold: 0.0024782463811330903  
Episode: 298, Score: -118.62338948773676, Threshold: 0.0024286814535104282  
Episode: 299, Score: -127.18046179295924, Threshold: 0.0023801078244402194

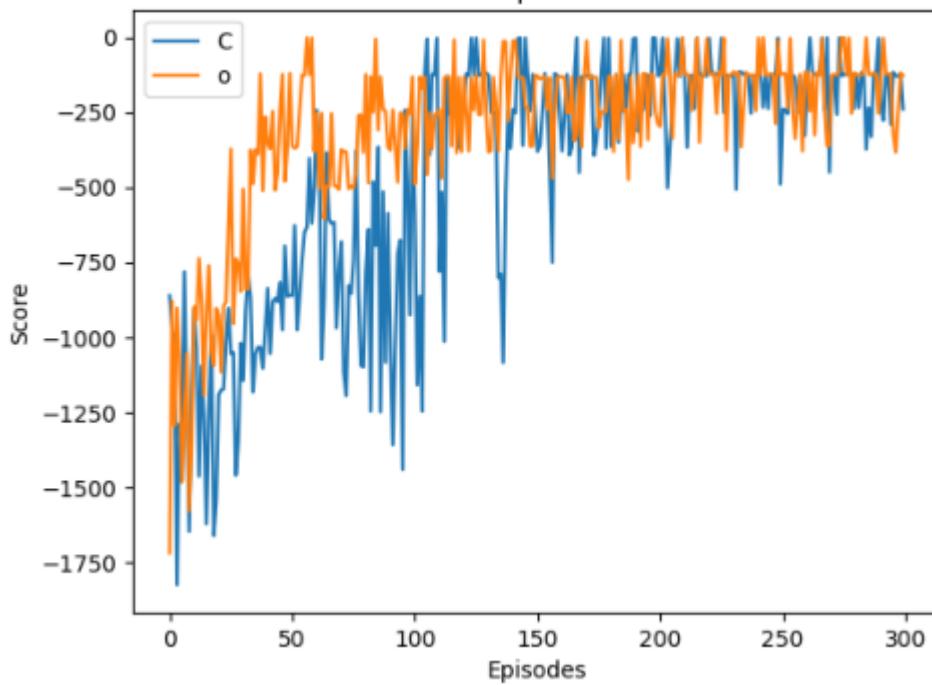


Comparison



```
In [94]: torch.save(DDAgent.policy_net.state_dict(), "./rl_models/double_dqn.pt")
showImage("./Img/RL/ddqn.png")
```

Comparison



```
In [96]: evaluate_model(DDAgent, 100)
```

Evaluating Episode: 1, Score: -132.44311160159268  
Evaluating Episode: 2, Score: -133.848964565773  
Evaluating Episode: 3, Score: -123.43368971899999  
Evaluating Episode: 4, Score: -251.41484798415826  
Evaluating Episode: 5, Score: -129.9899541733548  
Evaluating Episode: 6, Score: -8.725552251734054  
Evaluating Episode: 7, Score: -132.89988956976507  
Evaluating Episode: 8, Score: -242.72683304932437  
Evaluating Episode: 9, Score: -133.3335926104606  
Evaluating Episode: 10, Score: -366.9640009047146  
Evaluating Episode: 11, Score: -8.742260311312142  
Evaluating Episode: 12, Score: -133.8247788899434  
Evaluating Episode: 13, Score: -133.55853040907573  
Evaluating Episode: 14, Score: -279.7122677375049  
Evaluating Episode: 15, Score: -10.769411274796447  
Evaluating Episode: 16, Score: -252.79057778831626  
Evaluating Episode: 17, Score: -133.09159953673912  
Evaluating Episode: 18, Score: -11.359946725670483  
Evaluating Episode: 19, Score: -132.52267161434418  
Evaluating Episode: 20, Score: -129.6967052921864  
Evaluating Episode: 21, Score: -121.8944933425897  
Evaluating Episode: 22, Score: -251.9420767071638  
Evaluating Episode: 23, Score: -259.2405743523838  
Evaluating Episode: 24, Score: -241.49226382074724  
Evaluating Episode: 25, Score: -131.6690288897437  
Evaluating Episode: 26, Score: -9.179513301903976  
Evaluating Episode: 27, Score: -259.07057329841587  
Evaluating Episode: 28, Score: -336.8626548607074  
Evaluating Episode: 29, Score: -236.95692945163827  
Evaluating Episode: 30, Score: -124.92470942695215  
Evaluating Episode: 31, Score: -8.793251212847439  
Evaluating Episode: 32, Score: -125.47200007747465  
Evaluating Episode: 33, Score: -8.730015242138236  
Evaluating Episode: 34, Score: -257.10848663649534  
Evaluating Episode: 35, Score: -129.08424820114925  
Evaluating Episode: 36, Score: -135.1129410151572  
Evaluating Episode: 37, Score: -133.28397330499703  
Evaluating Episode: 38, Score: -8.75555744263936  
Evaluating Episode: 39, Score: -139.25931615558673  
Evaluating Episode: 40, Score: -360.0774364207425  
Evaluating Episode: 41, Score: -360.5966178558563  
Evaluating Episode: 42, Score: -127.55750422606863  
Evaluating Episode: 43, Score: -231.4811108604707  
Evaluating Episode: 44, Score: -133.69712816469223  
Evaluating Episode: 45, Score: -131.83153399612652  
Evaluating Episode: 46, Score: -9.193322122858259  
Evaluating Episode: 47, Score: -134.0991489339505  
Evaluating Episode: 48, Score: -134.5388763256594  
Evaluating Episode: 49, Score: -127.61905256130942  
Evaluating Episode: 50, Score: -133.68632537504553  
Evaluating Episode: 51, Score: -134.39536478607891  
Evaluating Episode: 52, Score: -241.0890255733878  
Evaluating Episode: 53, Score: -138.1892210568582  
Evaluating Episode: 54, Score: -9.06273322582651  
Evaluating Episode: 55, Score: -132.4080753412374  
Evaluating Episode: 56, Score: -121.3640445576526  
Evaluating Episode: 57, Score: -132.47724693969502  
Evaluating Episode: 58, Score: -133.32154320584814  
Evaluating Episode: 59, Score: -124.49605114081243  
Evaluating Episode: 60, Score: -10.316188842471796  
Evaluating Episode: 61, Score: -259.6624655190032  
Evaluating Episode: 62, Score: -132.54332586180988  
Evaluating Episode: 63, Score: -137.30148137181664  
Evaluating Episode: 64, Score: -129.73934708444708  
Evaluating Episode: 65, Score: -11.542227250968986  
Evaluating Episode: 66, Score: -259.01139742806856

```
Evaluating Episode: 67, Score: -132.69775473899335
Evaluating Episode: 68, Score: -130.56242760776053
Evaluating Episode: 69, Score: -8.960601790904619
Evaluating Episode: 70, Score: -131.93223573175896
Evaluating Episode: 71, Score: -125.16928026283404
Evaluating Episode: 72, Score: -127.26039464977767
Evaluating Episode: 73, Score: -9.295821704707071
Evaluating Episode: 74, Score: -252.1710398804282
Evaluating Episode: 75, Score: -9.670627593644472
Evaluating Episode: 76, Score: -388.75339967007807
Evaluating Episode: 77, Score: -267.6984479369628
Evaluating Episode: 78, Score: -129.84358766040918
Evaluating Episode: 79, Score: -129.36576982609444
Evaluating Episode: 80, Score: -236.71517372310015
Evaluating Episode: 81, Score: -131.3669469093519
Evaluating Episode: 82, Score: -244.91785633418257
Evaluating Episode: 83, Score: -128.23162905099122
Evaluating Episode: 84, Score: -131.80462411842103
Evaluating Episode: 85, Score: -128.89396678392526
Evaluating Episode: 86, Score: -124.08083924956621
Evaluating Episode: 87, Score: -8.994535449490604
Evaluating Episode: 88, Score: -249.38525223982833
Evaluating Episode: 89, Score: -128.9789775318166
Evaluating Episode: 90, Score: -376.96664986236675
Evaluating Episode: 91, Score: -124.98527321256886
Evaluating Episode: 92, Score: -274.07542240772807
Evaluating Episode: 93, Score: -126.81215185952387
Evaluating Episode: 94, Score: -121.59323876478003
Evaluating Episode: 95, Score: -132.38670254584287
Evaluating Episode: 96, Score: -136.30570298467163
Evaluating Episode: 97, Score: -249.85606576813964
Evaluating Episode: 98, Score: -237.6213096465926
Evaluating Episode: 99, Score: -10.312292494030565
Evaluating Episode: 100, Score: -134.26386811342817
Final Score: -150.73909456883962
```

```
In [97]: visualize_model(DDAgent,3)
```

```
Visualizing Episode: 1, Score: -8.456244914503419
Visualizing Episode: 2, Score: -8.60563807086996
Visualizing Episode: 3, Score: -136.2381554988296
```

### Observation:

### Stability:

This model is definitely less prone to sudden lapses. The accuracy rarely deteriorates although it oscillates from -250 to 0. This is slightly concerning but i think its not a deal breaker.

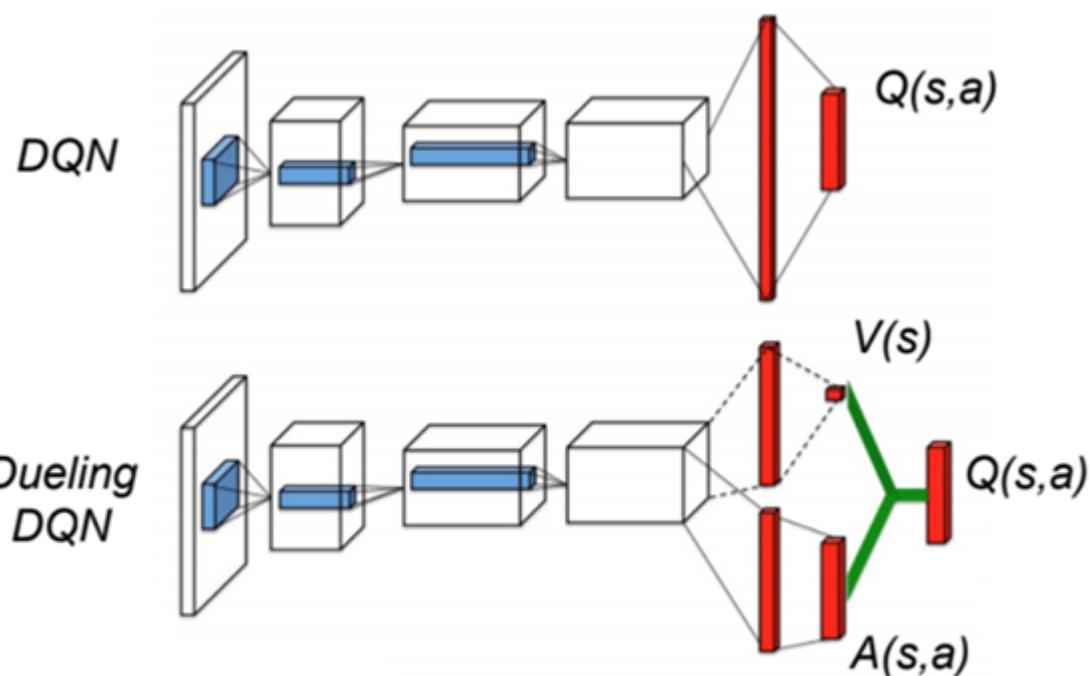
### Training Speed:

One of the biggest improvements is training time. Compared to the previous DQN which got to good accuracy after 100 episodes, this model reached that in around 60 episodes.

### Performance:

With a 100 epoch average of -150, we are literally the #7 best model for this Pendulum Problem according to the github leaderboard shown earlier. This is a significant improvement from the -160 I got earlier

## Double Dueling DQN



Dueling architectures aim to address the limitations of DDQN by separating the estimation of the value and advantage functions, providing more stable and accurate Q-value predictions.

### 1. Reducing Overestimation Bias:

- The dueling architecture helps mitigate overestimation bias by separating the estimation of state values and action advantages.

### 2. Enhanced Stability in Training:

- D3QN tends to exhibit improved stability during training, reducing the risk of oscillations in learning.

This can easily be implemented by modifying the MLP model.

```
In [98]: class DDQN_MLP(nn.Module):
    def __init__(self, inp, out, lr):
        super(DDQN_MLP, self).__init__()

        self.layer1 = nn.Linear(inp, 32)
        self.layer2 = nn.Linear(32, 64)

        self.value_stream = nn.Sequential(
            nn.Linear(64, 64),
            nn.LeakyReLU(inplace=True),
            nn.Linear(64, 1)
        )

        self.advantage_stream = nn.Sequential(
            nn.Linear(64, 64),
            nn.LeakyReLU(inplace=True),
            nn.Linear(64, out)
        )

        self.lr = lr
        self.optimizer = optim.Adam(self.parameters(), lr=self.lr)

    def forward(self, x):
        x = self.layer1(x)
        x = F.leaky_relu(x)
        x = self.layer2(x)
```

```

x = F.leaky_relu(x)
# x = self.layer3(x)
# x = F.leaky_relu(x)

values = self.value_stream(x)
advantages = self.advantage_stream(x)
q_values = values + advantages - advantages.mean()

return q_values

```

In [100]:

```

class D3DQNAgent:
    def __init__(self):
        self.steps = 0
        self.memory = Memory(10000)

        # Greedy Epsilon Hyperparams
        self.epsilon = 1.0
        self.epsilon_decay = 0.98
        self.epsilon_min = 0.001

        self.batch_size = 256
        self.adamlr = 0.001
        self.targetlr = 0.01
        self.gamma = 0.98

        self.numOfAction = 20
        self.action_array = np.concatenate((-1 * (np.exp(np.linspace(0, 1.09, 10)) - 1),))

        self.policy_net = DDQN_MLP(3, self.numOfAction, self.adamlr)
        self.target_net = DDQN_MLP(3, self.numOfAction, self.adamlr)
        self.target_net.load_state_dict(self.policy_net.state_dict())

    def transform_choice(self, num):
        return self.action_array[int(num)]

    def action(self, state):
        sample = np.random.rand()

        if sample > self.epsilon:
            with torch.no_grad():
                choice = float(torch.argmax(self.policy_net(state)).numpy())
                action = self.transform_choice(choice)
        else:
            choice = np.random.choice([n for n in range(self.numOfAction)])
            action = self.transform_choice(choice)

        return choice, action

    def optimize(self):
        if len(self.memory) < self.batch_size:
            return

        batch = self.memory.sample(self.batch_size)
        state, action, reward, next_state = batch
        action = action.type(torch.int64)

        with torch.no_grad():
            expected_q = self.target_net(next_state).max(1)[0].unsqueeze(1)
            expected = reward + self.gamma * expected_q

        # Actual values - Policy Net
        actual = self.policy_net(state).gather(1,action)

        # Back_propagation

```

```
loss = F.smooth_l1_loss(actual, expected)
self.policy_net.optimizer.zero_grad()
loss.mean().backward()
nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
self.policy_net.optimizer.step()

for target_param, param in zip(self.target_net.parameters(), self.policy_net.pa
    target_param.data.copy_(target_param.data * (1.0 - self.targetlr) + param
```

```
In [101...]: d3Agent, d3qn_score = train_model_eps(D3DQNAgent, 300)
plot_tuples([(ddqn_score, "base"), (d3qn_score, "action")], "Compare Dueling DQN")
```

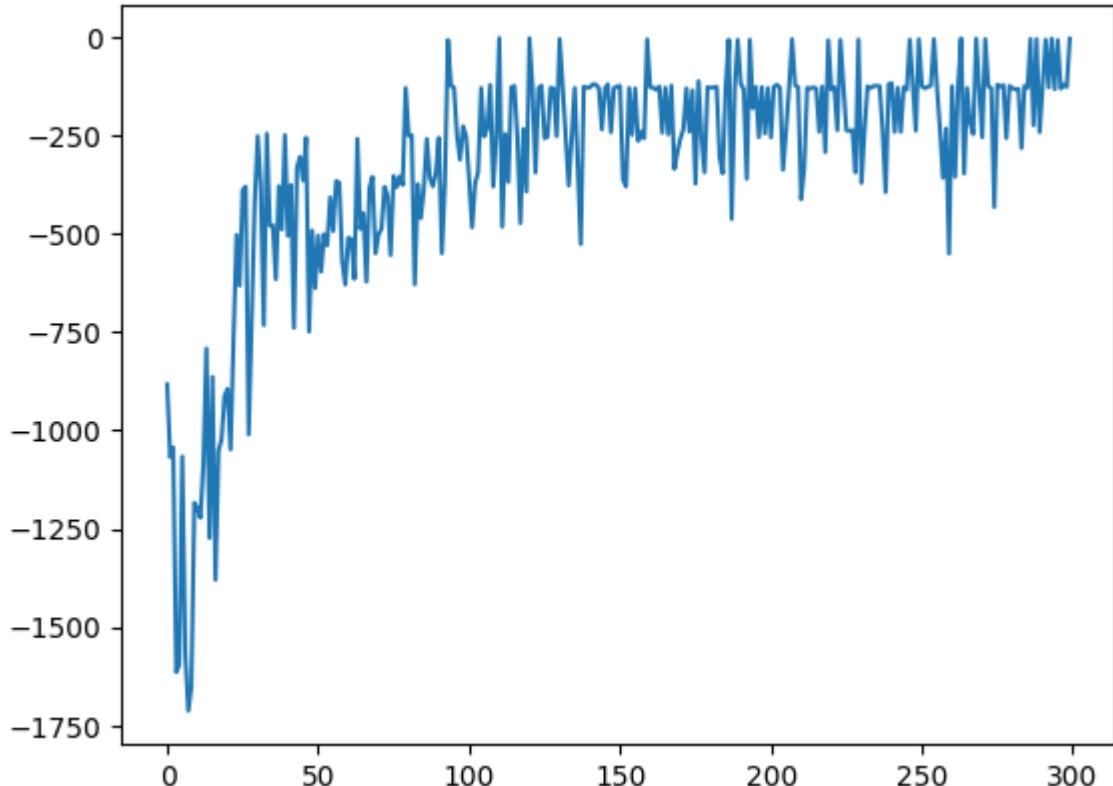
Episode: 0, Score: -882.9720168590786, Threshold: 1.0  
Episode: 1, Score: -1068.1073584916733, Threshold: 0.98  
Episode: 2, Score: -1043.747252454962, Threshold: 0.9603999999999999  
Episode: 3, Score: -1614.8689938988646, Threshold: 0.9411919999999999  
Episode: 4, Score: -1596.900833310135, Threshold: 0.9223681599999999  
Episode: 5, Score: -1066.2255004437843, Threshold: 0.9039207967999998  
Episode: 6, Score: -1568.8463418482115, Threshold: 0.8858423808639998  
Episode: 7, Score: -1713.1005602422458, Threshold: 0.8681255332467198  
Episode: 8, Score: -1653.0370023003052, Threshold: 0.8507630225817854  
Episode: 9, Score: -1184.1966162149529, Threshold: 0.8337477621301497  
Episode: 10, Score: -1197.8968483538235, Threshold: 0.8170728068875467  
Episode: 11, Score: -1222.4327573430426, Threshold: 0.8007313507497957  
Episode: 12, Score: -1083.775666196598, Threshold: 0.7847167237347998  
Episode: 13, Score: -791.8370683951285, Threshold: 0.7690223892601038  
Episode: 14, Score: -1273.95864899669, Threshold: 0.7536419414749017  
Episode: 15, Score: -863.7952267257812, Threshold: 0.7385691026454037  
Episode: 16, Score: -1380.1309974217722, Threshold: 0.7237977205924956  
Episode: 17, Score: -1051.1079822110923, Threshold: 0.7093217661806457  
Episode: 18, Score: -1023.7394980074866, Threshold: 0.6951353308570327  
Episode: 19, Score: -912.812992630354, Threshold: 0.6812326242398921  
Episode: 20, Score: -894.4487149638567, Threshold: 0.6676079717550942  
Episode: 21, Score: -1048.4615161624986, Threshold: 0.6542558123199923  
Episode: 22, Score: -760.3607674842425, Threshold: 0.6411706960735924  
Episode: 23, Score: -502.4486229664367, Threshold: 0.6283472821521205  
Episode: 24, Score: -632.0473837603671, Threshold: 0.6157803365090782  
Episode: 25, Score: -388.6753096646986, Threshold: 0.6034647297788965  
Episode: 26, Score: -380.12021599010956, Threshold: 0.5913954351833186  
Episode: 27, Score: -1010.3611745891053, Threshold: 0.5795675264796523  
Episode: 28, Score: -720.910316544517, Threshold: 0.5679761759500592  
Episode: 29, Score: -418.5790580410458, Threshold: 0.5566166524310581  
Episode: 30, Score: -251.4732428913829, Threshold: 0.5454843193824369  
Episode: 31, Score: -376.3165204927492, Threshold: 0.5345746329947881  
Episode: 32, Score: -732.0473633886388, Threshold: 0.5238831403348924  
Episode: 33, Score: -245.65713411944932, Threshold: 0.5134054775281945  
Episode: 34, Score: -479.40610446606337, Threshold: 0.5031373679776306  
Episode: 35, Score: -478.43570621374414, Threshold: 0.493074620618078  
Episode: 36, Score: -615.5947930948005, Threshold: 0.48321312820571644  
Episode: 37, Score: -378.49884813040853, Threshold: 0.4735488656416021  
Episode: 38, Score: -488.8966484774073, Threshold: 0.46407788832877006  
Episode: 39, Score: -249.1219698060636, Threshold: 0.45479633056219465  
Episode: 40, Score: -505.0478882240228, Threshold: 0.44570040395095073  
Episode: 41, Score: -375.7778498424019, Threshold: 0.4367863958719317  
Episode: 42, Score: -739.0327994140589, Threshold: 0.42805066795449304  
Episode: 43, Score: -331.4067112172851, Threshold: 0.41948965459540316  
Episode: 44, Score: -304.2627326358843, Threshold: 0.4110998615034951  
Episode: 45, Score: -364.29170779537253, Threshold: 0.4028778642734252  
Episode: 46, Score: -255.60483314429618, Threshold: 0.39482030698795667  
Episode: 47, Score: -748.7647120571427, Threshold: 0.38692390084819756  
Episode: 48, Score: -491.3539046216351, Threshold: 0.3791854228312336  
Episode: 49, Score: -638.2714449210189, Threshold: 0.37160171437460887  
Episode: 50, Score: -504.5689718574513, Threshold: 0.3641696800871167  
Episode: 51, Score: -595.9430892356467, Threshold: 0.35688628648537435  
Episode: 52, Score: -500.8607710798313, Threshold: 0.34974856075566685  
Episode: 53, Score: -529.5600874206863, Threshold: 0.3427535895405535  
Episode: 54, Score: -407.0977619205881, Threshold: 0.33589851774974244  
Episode: 55, Score: -492.81398273487093, Threshold: 0.3291805473947476  
Episode: 56, Score: -365.8262188937034, Threshold: 0.32259693644685267  
Episode: 57, Score: -370.3820946887004, Threshold: 0.3161449977179156  
Episode: 58, Score: -569.6244572780292, Threshold: 0.3098220977635573  
Episode: 59, Score: -629.093172555378, Threshold: 0.30362565580828615  
Episode: 60, Score: -510.1442598398063, Threshold: 0.2975531426921204  
Episode: 61, Score: -514.0280781424992, Threshold: 0.291602079838278  
Episode: 62, Score: -614.5237823976998, Threshold: 0.2857700382415124  
Episode: 63, Score: -258.48835541769205, Threshold: 0.2800546374766822  
Episode: 64, Score: -487.40889183754405, Threshold: 0.27445354472714856  
Episode: 65, Score: -448.0107322394201, Threshold: 0.2689644738326056

Episode: 66, Score: -621.895830844793, Threshold: 0.26358518435595346  
Episode: 67, Score: -383.54902663197726, Threshold: 0.25831348066883436  
Episode: 68, Score: -355.0569986759363, Threshold: 0.2531472110554577  
Episode: 69, Score: -549.4554512867361, Threshold: 0.24808426683434853  
Episode: 70, Score: -502.3741982645566, Threshold: 0.24312258149766156  
Episode: 71, Score: -486.84355186577403, Threshold: 0.2382601298677083  
Episode: 72, Score: -380.37068833875236, Threshold: 0.23349492727035415  
Episode: 73, Score: -403.1561614013442, Threshold: 0.22882502872494706  
Episode: 74, Score: -554.1098260322793, Threshold: 0.22424852815044813  
Episode: 75, Score: -352.94024657250105, Threshold: 0.21976355758743915  
Episode: 76, Score: -381.55427997195795, Threshold: 0.21536828643569036  
Episode: 77, Score: -353.88849927162204, Threshold: 0.21106092070697655  
Episode: 78, Score: -376.0284554491783, Threshold: 0.20683970229283702  
Episode: 79, Score: -129.09260728103172, Threshold: 0.20270290824698028  
Episode: 80, Score: -250.4415235680376, Threshold: 0.19864885008204067  
Episode: 81, Score: -248.9787414226223, Threshold: 0.19467587308039985  
Episode: 82, Score: -628.9015520954824, Threshold: 0.19078235561879187  
Episode: 83, Score: -373.5644096602649, Threshold: 0.18696670850641603  
Episode: 84, Score: -460.3403047993519, Threshold: 0.18322737433628772  
Episode: 85, Score: -388.1238324701789, Threshold: 0.17956282684956196  
Episode: 86, Score: -258.62381417636226, Threshold: 0.1759715703125707  
Episode: 87, Score: -357.967815225389, Threshold: 0.1724521389063193  
Episode: 88, Score: -379.4556575687268, Threshold: 0.16900309612819292  
Episode: 89, Score: -344.58336820623293, Threshold: 0.16562303420562907  
Episode: 90, Score: -255.84402578325378, Threshold: 0.16231057352151648  
Episode: 91, Score: -549.5037337934552, Threshold: 0.15906436205108615  
Episode: 92, Score: -366.5790028357161, Threshold: 0.15588307481006441  
Episode: 93, Score: -7.176996669563573, Threshold: 0.15276541331386312  
Episode: 94, Score: -123.67321617085912, Threshold: 0.14971010504758586  
Episode: 95, Score: -127.86174343027035, Threshold: 0.14671590294663414  
Episode: 96, Score: -249.46704112698455, Threshold: 0.14378158488770146  
Episode: 97, Score: -311.6624016084893, Threshold: 0.14090595318994742  
Episode: 98, Score: -226.63821128931494, Threshold: 0.13808783412614847  
Episode: 99, Score: -251.02046024058356, Threshold: 0.13532607744362551  
Episode: 100, Score: -359.2718035823477, Threshold: 0.132619555894753  
Episode: 101, Score: -483.52144608971213, Threshold: 0.12996716477685794  
Episode: 102, Score: -371.5551108333519, Threshold: 0.12736782148132078  
Episode: 103, Score: -342.95049544301094, Threshold: 0.12482046505169436  
Episode: 104, Score: -129.0982299218295, Threshold: 0.12232405575066048  
Episode: 105, Score: -252.58225157180306, Threshold: 0.11987757463564727  
Episode: 106, Score: -232.11125229649014, Threshold: 0.11748002314293432  
Episode: 107, Score: -121.39226948809204, Threshold: 0.11513042268007563  
Episode: 108, Score: -379.6493313319833, Threshold: 0.11282781422647412  
Episode: 109, Score: -253.2831810694056, Threshold: 0.11057125794194463  
Episode: 110, Score: -2.1323453536226302, Threshold: 0.10835983278310574  
Episode: 111, Score: -481.9358959721466, Threshold: 0.10619263612744362  
Episode: 112, Score: -246.14855695912388, Threshold: 0.10406878340489474  
Episode: 113, Score: -368.03038386137166, Threshold: 0.10198740773679685  
Episode: 114, Score: -127.61030131252751, Threshold: 0.09994765958206091  
Episode: 115, Score: -124.0365159024369, Threshold: 0.0979487063904197  
Episode: 116, Score: -255.78785136141985, Threshold: 0.0959897322626113  
Episode: 117, Score: -472.8977740398889, Threshold: 0.09406993761735907  
Episode: 118, Score: -233.4307794362701, Threshold: 0.0921885388650119  
Episode: 119, Score: -392.21957150863926, Threshold: 0.09034476808771166  
Episode: 120, Score: -2.7117035929113538, Threshold: 0.08853787272595742  
Episode: 121, Score: -127.28692202098308, Threshold: 0.08676711527143827  
Episode: 122, Score: -344.2909496711081, Threshold: 0.0850317729660095  
Episode: 123, Score: -129.05157978848578, Threshold: 0.08333113750668932  
Episode: 124, Score: -123.44282040404892, Threshold: 0.08166451475655553  
Episode: 125, Score: -257.0279163835628, Threshold: 0.08003122446142442  
Episode: 126, Score: -253.402865058023, Threshold: 0.07843059997219594  
Episode: 127, Score: -128.5628057969308, Threshold: 0.07686198797275202  
Episode: 128, Score: -130.18076896110568, Threshold: 0.07532474821329699  
Episode: 129, Score: -251.6903350694914, Threshold: 0.07381825324903105  
Episode: 130, Score: -4.69415050691886, Threshold: 0.07234188818405042  
Episode: 131, Score: -125.14652197323022, Threshold: 0.07089505042036941

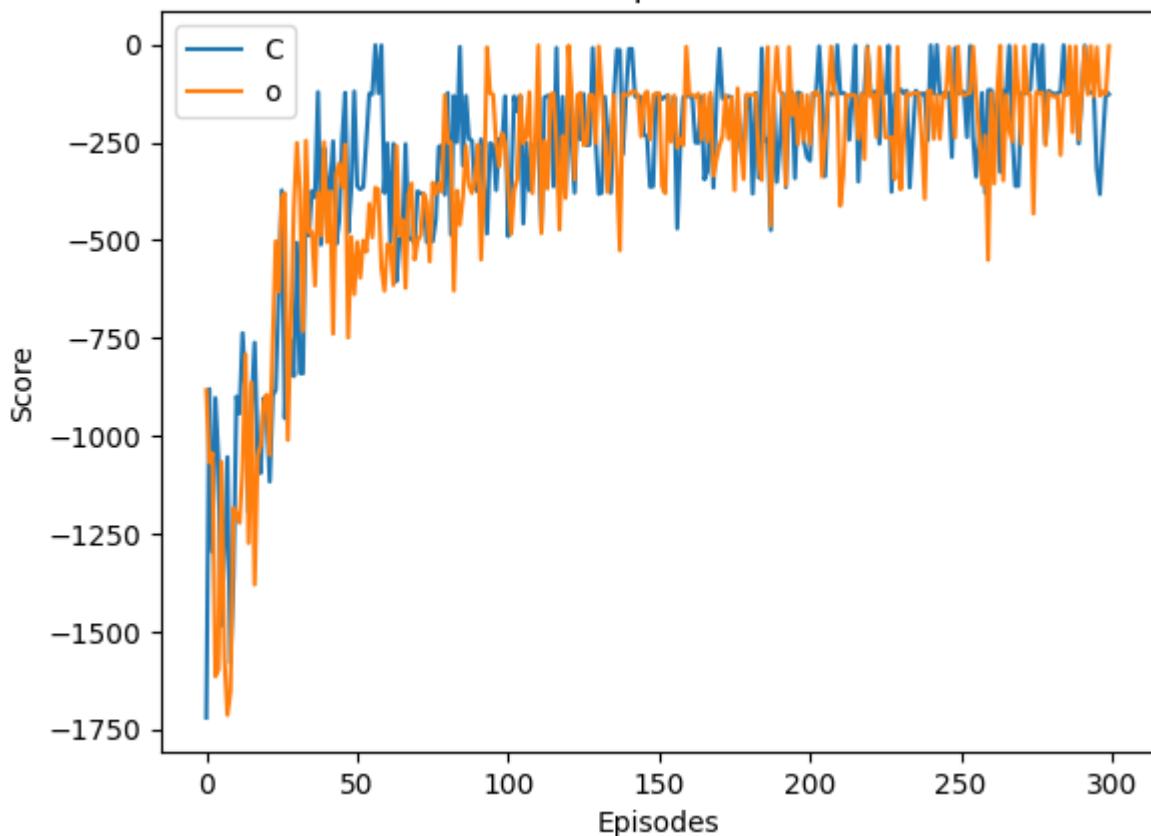
Episode: 132, Score: -254.5852475181984, Threshold: 0.06947714941196202  
Episode: 133, Score: -377.26029153953465, Threshold: 0.06808760642372277  
Episode: 134, Score: -254.92729062896652, Threshold: 0.06672585429524831  
Episode: 135, Score: -129.22867272038556, Threshold: 0.06539133720934334  
Episode: 136, Score: -326.8146706629547, Threshold: 0.06408351046515648  
Episode: 137, Score: -526.1009638526094, Threshold: 0.06280184025585335  
Episode: 138, Score: -125.66720336582571, Threshold: 0.061545803450736285  
Episode: 139, Score: -128.70909493320372, Threshold: 0.060314887381721555  
Episode: 140, Score: -127.79578725570444, Threshold: 0.05910858963408712  
Episode: 141, Score: -119.6506597331556, Threshold: 0.05792641784140538  
Episode: 142, Score: -120.46678248151474, Threshold: 0.05676788948457727  
Episode: 143, Score: -129.47686787452926, Threshold: 0.055632531694885724  
Episode: 144, Score: -234.6358091228506, Threshold: 0.054519881060988006  
Episode: 145, Score: -130.78051487360202, Threshold: 0.05342948343976824  
Episode: 146, Score: -119.82877335354087, Threshold: 0.052360893770972874  
Episode: 147, Score: -241.8517785876699, Threshold: 0.05131367589555342  
Episode: 148, Score: -129.45173951854105, Threshold: 0.05028740237764235  
Episode: 149, Score: -125.56600608907199, Threshold: 0.0492816543300895  
Episode: 150, Score: -124.06162171488832, Threshold: 0.04829602124348771  
Episode: 151, Score: -361.5994414235859, Threshold: 0.04733010081861796  
Episode: 152, Score: -379.95450106802997, Threshold: 0.0463834988022456  
Episode: 153, Score: -129.1493480278569, Threshold: 0.04545582882620069  
Episode: 154, Score: -248.25894025086166, Threshold: 0.04454671224967667  
Episode: 155, Score: -130.10422542333444, Threshold: 0.043655778004683135  
Episode: 156, Score: -263.74314998777345, Threshold: 0.04278266244458947  
Episode: 157, Score: -240.4182320213822, Threshold: 0.04192700919569768  
Episode: 158, Score: -257.0598915505443, Threshold: 0.04108846901178373  
Episode: 159, Score: -5.657707196204955, Threshold: 0.04026669963154805  
Episode: 160, Score: -127.33434446448051, Threshold: 0.03946136563891709  
Episode: 161, Score: -128.4359330204466, Threshold: 0.03867213832613875  
Episode: 162, Score: -133.83809169597959, Threshold: 0.037898695559615975  
Episode: 163, Score: -125.19274575080574, Threshold: 0.03714072164842366  
Episode: 164, Score: -242.80902276232393, Threshold: 0.03639790721545518  
Episode: 165, Score: -128.29278357289542, Threshold: 0.035669949071146075  
Episode: 166, Score: -247.1818511476569, Threshold: 0.034956550089723155  
Episode: 167, Score: -123.21471371201338, Threshold: 0.03425741908792869  
Episode: 168, Score: -334.91111620749996, Threshold: 0.03357227070617012  
Episode: 169, Score: -293.0762218939253, Threshold: 0.032900825292046715  
Episode: 170, Score: -257.9016398249978, Threshold: 0.03224280878620578  
Episode: 171, Score: -234.1986118134012, Threshold: 0.031597952610481664  
Episode: 172, Score: -129.93593300255114, Threshold: 0.03096599355827203  
Episode: 173, Score: -240.54004611959527, Threshold: 0.030346673687106588  
Episode: 174, Score: -135.03729340900767, Threshold: 0.029739740213364455  
Episode: 175, Score: -372.40272450376426, Threshold: 0.029144945409097164  
Episode: 176, Score: -111.81526864575855, Threshold: 0.02856204650091522  
Episode: 177, Score: -255.11025185398574, Threshold: 0.027990805570896914  
Episode: 178, Score: -343.68724898004245, Threshold: 0.027430989459478975  
Episode: 179, Score: -126.86038048413727, Threshold: 0.026882369670289397  
Episode: 180, Score: -128.55796377471208, Threshold: 0.026344722276883607  
Episode: 181, Score: -127.91370164249871, Threshold: 0.025817827831345935  
Episode: 182, Score: -125.97936466604293, Threshold: 0.025301471274719014  
Episode: 183, Score: -304.4306632832539, Threshold: 0.024795441849224633  
Episode: 184, Score: -346.4667334413771, Threshold: 0.02429953301224014  
Episode: 185, Score: -131.28586892741384, Threshold: 0.02381354235199534  
Episode: 186, Score: -6.566194339254657, Threshold: 0.02333727150495543  
Episode: 187, Score: -462.17918913298297, Threshold: 0.02287052607485632  
Episode: 188, Score: -130.8281948047199, Threshold: 0.022413115553359196  
Episode: 189, Score: -6.687573878563092, Threshold: 0.021964853242292012  
Episode: 190, Score: -118.13073689410051, Threshold: 0.02152555617744617  
Episode: 191, Score: -133.74684100383098, Threshold: 0.021095045053897248  
Episode: 192, Score: -360.0468594479798, Threshold: 0.020673144152819303  
Episode: 193, Score: -6.64083446462541, Threshold: 0.02025968126976292  
Episode: 194, Score: -180.49659615531183, Threshold: 0.01985448764436766  
Episode: 195, Score: -126.88688730280285, Threshold: 0.019457397891480308  
Episode: 196, Score: -254.41046024128005, Threshold: 0.0190682499336507  
Episode: 197, Score: -126.9933025966254, Threshold: 0.018686884934977686

Episode: 198, Score: -245.04844206844757, Threshold: 0.01831314723627813  
Episode: 199, Score: -129.39030964876284, Threshold: 0.017946884291552567  
Episode: 200, Score: -254.90510206979494, Threshold: 0.017587946605721515  
Episode: 201, Score: -125.8122846178257, Threshold: 0.017236187673607083  
Episode: 202, Score: -120.83918877741513, Threshold: 0.016891463920134943  
Episode: 203, Score: -129.05830541306537, Threshold: 0.016553634641732245  
Episode: 204, Score: -337.2813241325111, Threshold: 0.0162225619488976  
Episode: 205, Score: -238.48415697045726, Threshold: 0.015898110709919647  
Episode: 206, Score: -130.48509905654487, Threshold: 0.015580148495721254  
Episode: 207, Score: -5.020879903796819, Threshold: 0.015268545525806828  
Episode: 208, Score: -122.58902017387135, Threshold: 0.014963174615290691  
Episode: 209, Score: -126.88314807443388, Threshold: 0.014663911122984877  
Episode: 210, Score: -412.0424916363688, Threshold: 0.01437063290052518  
Episode: 211, Score: -334.4298296547298, Threshold: 0.014083220242514675  
Episode: 212, Score: -128.7703298110943, Threshold: 0.013801555837664382  
Episode: 213, Score: -130.76333591269213, Threshold: 0.013525524720911095  
Episode: 214, Score: -127.76272013064747, Threshold: 0.013255014226492874  
Episode: 215, Score: -130.5152377726961, Threshold: 0.012989913941963017  
Episode: 216, Score: -240.75970110077049, Threshold: 0.012730115663123756  
Episode: 217, Score: -125.37429653547684, Threshold: 0.012475513349861281  
Episode: 218, Score: -292.5243810886951, Threshold: 0.012226003082864055  
Episode: 219, Score: -7.824805251162482, Threshold: 0.011981483021206774  
Episode: 220, Score: -132.81745526122316, Threshold: 0.011741853360782638  
Episode: 221, Score: -127.53539155204469, Threshold: 0.011507016293566985  
Episode: 222, Score: -236.54259806971277, Threshold: 0.011276875967695645  
Episode: 223, Score: -6.865285260580511, Threshold: 0.011051338448341731  
Episode: 224, Score: -123.43184619774007, Threshold: 0.010830311679374895  
Episode: 225, Score: -236.9768985106987, Threshold: 0.010613705445787398  
Episode: 226, Score: -237.89120991316534, Threshold: 0.01040143133687165  
Episode: 227, Score: -236.56880291202432, Threshold: 0.010193402710134216  
Episode: 228, Score: -343.84401590903565, Threshold: 0.009989534655931531  
Episode: 229, Score: -6.123921543073717, Threshold: 0.0097897439628129  
Episode: 230, Score: -370.4083673692726, Threshold: 0.009593949083556642  
Episode: 231, Score: -237.46403003288845, Threshold: 0.009402070101885509  
Episode: 232, Score: -125.00130479876829, Threshold: 0.009214028699847799  
Episode: 233, Score: -129.23894319687844, Threshold: 0.009029748125850843  
Episode: 234, Score: -124.30253896149685, Threshold: 0.008849153163333826  
Episode: 235, Score: -123.77149117086557, Threshold: 0.008672170100067149  
Episode: 236, Score: -124.48227619556451, Threshold: 0.008498726698065806  
Episode: 237, Score: -236.7356736422591, Threshold: 0.00832875216410449  
Episode: 238, Score: -394.2000279513022, Threshold: 0.0081621771208224  
Episode: 239, Score: -119.36951274810806, Threshold: 0.007998933578405951  
Episode: 240, Score: -117.04499697506054, Threshold: 0.007838954906837833  
Episode: 241, Score: -240.83977213949052, Threshold: 0.007682175808701076  
Episode: 242, Score: -128.50668175139953, Threshold: 0.007528532292527055  
Episode: 243, Score: -240.81088643207855, Threshold: 0.007377961646676513  
Episode: 244, Score: -127.24632883552077, Threshold: 0.007230402413742983  
Episode: 245, Score: -133.68880537149315, Threshold: 0.007085794365468123  
Episode: 246, Score: -6.426328282745229, Threshold: 0.00694407847815876  
Episode: 247, Score: -125.50515852859729, Threshold: 0.0068051969085955845  
Episode: 248, Score: -237.72507473630222, Threshold: 0.006669092970423673  
Episode: 249, Score: -5.226693797390046, Threshold: 0.006535711111015199  
Episode: 250, Score: -124.83826911053848, Threshold: 0.006404996888794895  
Episode: 251, Score: -130.05815103741563, Threshold: 0.006276896951018997  
Episode: 252, Score: -126.07224462582865, Threshold: 0.006151359011998617  
Episode: 253, Score: -123.13929042044026, Threshold: 0.006028331831758644  
Episode: 254, Score: -4.977942254735455, Threshold: 0.005907765195123471  
Episode: 255, Score: -126.6439513163639, Threshold: 0.005789609891221002  
Episode: 256, Score: -233.33143636533538, Threshold: 0.005673817693396582  
Episode: 257, Score: -357.637327768143, Threshold: 0.00556034133952865  
Episode: 258, Score: -232.46484826687686, Threshold: 0.005449134512738076  
Episode: 259, Score: -550.34768966105, Threshold: 0.0053401518224833145  
Episode: 260, Score: -124.3229262778388, Threshold: 0.005233348786033648  
Episode: 261, Score: -355.0985013304544, Threshold: 0.005128681810312975  
Episode: 262, Score: -125.38591415747734, Threshold: 0.005026108174106716  
Episode: 263, Score: -2.833734217706484, Threshold: 0.004925586010624581

Episode: 264, Score: -346.9634188670064, Threshold: 0.004827074290412089  
Episode: 265, Score: -129.5384267650911, Threshold: 0.004730532804603847  
Episode: 266, Score: -218.6455788772541, Threshold: 0.00463592214851177  
Episode: 267, Score: -246.37111641487036, Threshold: 0.004543203705541535  
Episode: 268, Score: -3.570597134335776, Threshold: 0.004452339631430704  
Episode: 269, Score: -125.18981654951172, Threshold: 0.00436329283880209  
Episode: 270, Score: -254.544148792524, Threshold: 0.004276026982026048  
Episode: 271, Score: -4.799869405542542, Threshold: 0.004190506442385527  
Episode: 272, Score: -124.98396171850179, Threshold: 0.004106696313537816  
Episode: 273, Score: -130.06392543752804, Threshold: 0.00402456238726706  
Episode: 274, Score: -431.808952068225, Threshold: 0.003944071139521719  
Episode: 275, Score: -120.75252610786498, Threshold: 0.0038651897167312848  
Episode: 276, Score: -125.23021541195445, Threshold: 0.003787885922396659  
Episode: 277, Score: -122.1199362161925, Threshold: 0.003712128203948726  
Episode: 278, Score: -256.6817293698642, Threshold: 0.0036378856398697514  
Episode: 279, Score: -123.91945921624054, Threshold: 0.0035651279270723563  
Episode: 280, Score: -129.45853709414652, Threshold: 0.0034938253685309092  
Episode: 281, Score: -133.0286045020712, Threshold: 0.003423948861160291  
Episode: 282, Score: -131.19386110553427, Threshold: 0.0033554698839370852  
Episode: 283, Score: -282.1157915762825, Threshold: 0.0032883604862583436  
Episode: 284, Score: -123.88236238512094, Threshold: 0.0032225932765331767  
Episode: 285, Score: -129.0762146855694, Threshold: 0.003158141411002513  
Episode: 286, Score: -4.406472349912652, Threshold: 0.003094978582782463  
Episode: 287, Score: -224.0571751987684, Threshold: 0.0030330790111268136  
Episode: 288, Score: -4.630508079124666, Threshold: 0.0029724174309042773  
Episode: 289, Score: -241.62632318654815, Threshold: 0.002912969082286192  
Episode: 290, Score: -120.56051705012683, Threshold: 0.002854709700640468  
Episode: 291, Score: -6.620679187502974, Threshold: 0.0027976155066276585  
Episode: 292, Score: -127.5508422045304, Threshold: 0.0027416631964951054  
Episode: 293, Score: -3.5949275600026067, Threshold: 0.0026868299325652034  
Episode: 294, Score: -132.57343994297838, Threshold: 0.002633093333913899  
Episode: 295, Score: -7.126901871611591, Threshold: 0.002580431467235621  
Episode: 296, Score: -130.76237563645333, Threshold: 0.0025288228378909086  
Episode: 297, Score: -119.92543608847048, Threshold: 0.0024782463811330903  
Episode: 298, Score: -126.43295055693277, Threshold: 0.0024286814535104282  
Episode: 299, Score: -3.7939314722656983, Threshold: 0.0023801078244402194

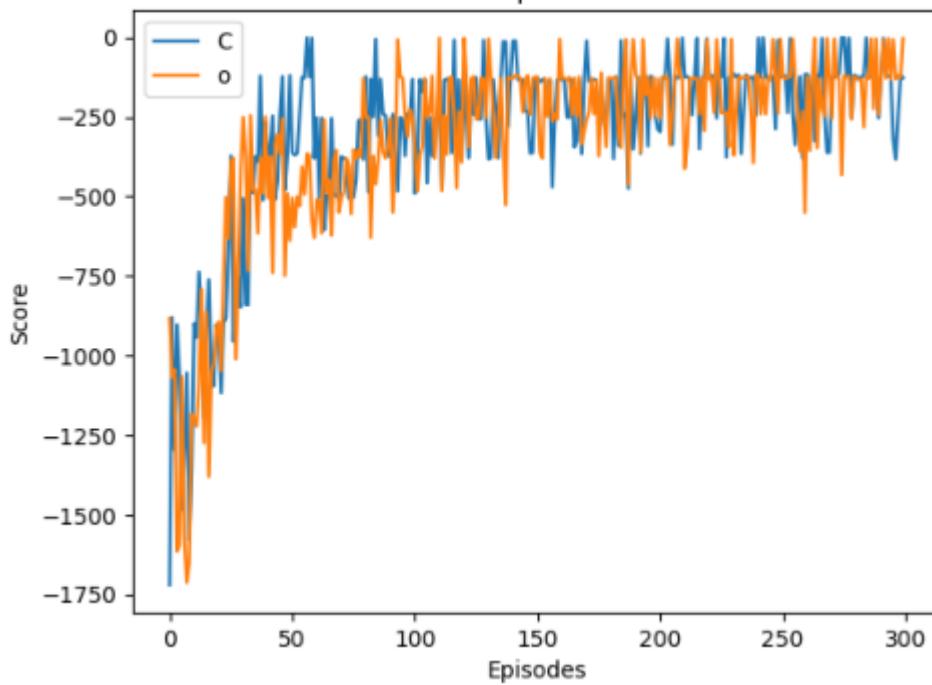


Comparison



```
In [102]: torch.save(d3Agent.policy_net.state_dict(), "./rl_models/d3_dqn.pt")
showImage("./Img/RL/d3qn.png")
```

Comparison



```
In [104]: evaluate_model(d3Agent, 100)
```

Evaluating Episode: 1, Score: -127.59256772621004  
Evaluating Episode: 2, Score: -239.81248402371136  
Evaluating Episode: 3, Score: -238.50409049633672  
Evaluating Episode: 4, Score: -259.6025446165501  
Evaluating Episode: 5, Score: -234.36524054949578  
Evaluating Episode: 6, Score: -242.11989743389043  
Evaluating Episode: 7, Score: -255.5806229434536  
Evaluating Episode: 8, Score: -235.01063982145246  
Evaluating Episode: 9, Score: -120.38471857947727  
Evaluating Episode: 10, Score: -118.96169287981972  
Evaluating Episode: 11, Score: -245.2257182042775  
Evaluating Episode: 12, Score: -1.608038043264248  
Evaluating Episode: 13, Score: -120.41615863309677  
Evaluating Episode: 14, Score: -247.67468755660587  
Evaluating Episode: 15, Score: -127.84812739998765  
Evaluating Episode: 16, Score: -129.12103517328725  
Evaluating Episode: 17, Score: -246.53165757065838  
Evaluating Episode: 18, Score: -2.1141059999520984  
Evaluating Episode: 19, Score: -362.9283969665194  
Evaluating Episode: 20, Score: -1.8854551369325065  
Evaluating Episode: 21, Score: -375.74061607839894  
Evaluating Episode: 22, Score: -115.14162999675648  
Evaluating Episode: 23, Score: -125.94995381142644  
Evaluating Episode: 24, Score: -124.18054772889302  
Evaluating Episode: 25, Score: -248.5713493011811  
Evaluating Episode: 26, Score: -117.96243035141902  
Evaluating Episode: 27, Score: -124.9787843105126  
Evaluating Episode: 28, Score: -2.539241766685915  
Evaluating Episode: 29, Score: -238.4536023654143  
Evaluating Episode: 30, Score: -263.2193732019183  
Evaluating Episode: 31, Score: -2.389291177247148  
Evaluating Episode: 32, Score: -361.4846302679067  
Evaluating Episode: 33, Score: -116.59832506979689  
Evaluating Episode: 34, Score: -125.46359376189405  
Evaluating Episode: 35, Score: -120.25877817262163  
Evaluating Episode: 36, Score: -129.09030825770327  
Evaluating Episode: 37, Score: -125.37407095474036  
Evaluating Episode: 38, Score: -120.14370050154841  
Evaluating Episode: 39, Score: -122.2980895305757  
Evaluating Episode: 40, Score: -118.71687237125435  
Evaluating Episode: 41, Score: -234.7045607232564  
Evaluating Episode: 42, Score: -128.19813983006017  
Evaluating Episode: 43, Score: -126.63725904268541  
Evaluating Episode: 44, Score: -1.9789204082428937  
Evaluating Episode: 45, Score: -238.3748538750089  
Evaluating Episode: 46, Score: -234.02322763279705  
Evaluating Episode: 47, Score: -224.03187466930305  
Evaluating Episode: 48, Score: -388.3521694038564  
Evaluating Episode: 49, Score: -115.75503711192351  
Evaluating Episode: 50, Score: -118.73783061410855  
Evaluating Episode: 51, Score: -247.66406792326137  
Evaluating Episode: 52, Score: -115.84955551027495  
Evaluating Episode: 53, Score: -1.7236906267829528  
Evaluating Episode: 54, Score: -230.14906406656988  
Evaluating Episode: 55, Score: -239.1500630413902  
Evaluating Episode: 56, Score: -120.4679612921205  
Evaluating Episode: 57, Score: -122.5184584787135  
Evaluating Episode: 58, Score: -126.02582403174297  
Evaluating Episode: 59, Score: -225.10618591642114  
Evaluating Episode: 60, Score: -127.91810198844229  
Evaluating Episode: 61, Score: -2.161866668398692  
Evaluating Episode: 62, Score: -346.47967474875674  
Evaluating Episode: 63, Score: -1.8316210979849987  
Evaluating Episode: 64, Score: -120.83364228330653  
Evaluating Episode: 65, Score: -222.4130361037713  
Evaluating Episode: 66, Score: -126.51270933429137

```
Evaluating Episode: 67, Score: -118.06603159660085
Evaluating Episode: 68, Score: -235.04198583598318
Evaluating Episode: 69, Score: -125.0543167427649
Evaluating Episode: 70, Score: -228.46787477979387
Evaluating Episode: 71, Score: -123.23367428233351
Evaluating Episode: 72, Score: -120.40674855734629
Evaluating Episode: 73, Score: -2.4165808739710943
Evaluating Episode: 74, Score: -118.5218349608685
Evaluating Episode: 75, Score: -247.01660053274333
Evaluating Episode: 76, Score: -2.4779378921888986
Evaluating Episode: 77, Score: -130.40815538648468
Evaluating Episode: 78, Score: -240.18695667630516
Evaluating Episode: 79, Score: -345.6203317127344
Evaluating Episode: 80, Score: -234.27333240188673
Evaluating Episode: 81, Score: -454.37690606225493
Evaluating Episode: 82, Score: -248.64813116195342
Evaluating Episode: 83, Score: -123.56151240011603
Evaluating Episode: 84, Score: -124.84567989743425
Evaluating Episode: 85, Score: -468.29101806626505
Evaluating Episode: 86, Score: -126.30156365538414
Evaluating Episode: 87, Score: -246.4409684057646
Evaluating Episode: 88, Score: -439.62104285215946
Evaluating Episode: 89, Score: -123.22687089773007
Evaluating Episode: 90, Score: -228.17380827592038
Evaluating Episode: 91, Score: -126.7939162690368
Evaluating Episode: 92, Score: -233.25106431077347
Evaluating Episode: 93, Score: -223.82144766454732
Evaluating Episode: 94, Score: -1.5536424208576494
Evaluating Episode: 95, Score: -121.6620458809439
Evaluating Episode: 96, Score: -236.23434887653409
Evaluating Episode: 97, Score: -129.11318920243428
Evaluating Episode: 98, Score: -125.47101873526412
Evaluating Episode: 99, Score: -3.3631228484738966
Evaluating Episode: 100, Score: -1.5720329908289754
Final Score: -168.80958158263027
```

Observations:

- Dueling Double DQN is nearly identical to Double DQN in terms of performances and stability.

While Dueling Double Deep Q-Network (D3QN) is a superior architecture, its impact might be limited when applied to the Pendulum problem. This is because:

### Complexity of Pendulum Dynamics:

- The Pendulum problem has relatively simple dynamics compared to more complex environments. The advantages of D3QN, such as reduced overestimation bias and enhanced stability, might not be as important when the underlying dynamics are straightforward.

### Limited Advantage Separation:

- In the Pendulum environment, the separation of value and advantage functions in the dueling architecture may not provide a significant advantage. The advantage stream might not effectively capture nuanced differences between actions, limiting the benefits of the dueling structure.

While D3QN is a powerful enhancement for certain reinforcement learning challenges, its impact may be limited in this case. Therefore, I would not use it further.

## Enhancing Memory Sampling

Efficient learning relies on effective exploration and utilization of past experiences stored in the replay buffer.

Improvement: In standard DQN training, experiences are uniformly sampled from the replay buffer, treating all transitions equally. I would like to change this by introducing Prioritized Replay which adopts a dynamic approach to memory sampling by assigning higher probabilities to experiences with larger TD errors.

Benefits:

1. **Focused Learning:** Prioritizing experiences with higher learning potential.
2. **Efficient Exploration:** Emphasizing crucial transitions for more effective training.

## Prioritized Experience Replay

Some experiences are more valuable than others. To take into account the importance of some experiences, we introduce prioritised experience replay.

How do we give a value to an experience?

- We give a higher value to experiences that lead to higher loss. We use the formula  $\text{abs}(\text{TD error}) + \text{constant}$
- A constant is added to the priority value so as to make sure some samples being given 0 priority

Stochastic Priorizations

- If we implement the above system without any modification, we would find only some experiences being used and some not at all.
- To mitigate this, we assign a probability to each experience this formula

$$P(i) = \frac{p_i^a}{\sum_k p_k^a}$$

Priority value

Normalized by all priority values in Replay Buffer

Hyperparameter used to reintroduce some randomness in the experience selection for the replay buffer

If  $a = 0$  pure uniform randomness

If  $a = 1$  only select the experiences with the highest priorities

Over here, "a" is a hyperparameter to be tuned.

```
In [109...]: class PERMemory(object):
```

```
def __init__(self, maxlen):  
    self.buffer = deque([], maxlen=maxlen)
```

```

    self.priorities = deque([], maxlen=maxlen)
    # using deque from collections but possible implement from scratch
def push(self, transition):
    self.buffer.append(transition)
    self.priorities.append(max(self.priorities, default=1))

def get_probs(self, alpha):
    #print(self.priorities)
    priorities = np.array(self.priorities) ** alpha
    sample_probs = priorities / np.sum(priorities)
    return sample_probs

def get_importance(self, probs):
    impt = 1/len(self.buffer) * 1/probs
    impt = impt / max(impt)
    return impt

def sample(self, batch_size, alpha=1):
    sample_probs = self.get_probs(alpha)
    samples_index = random.choices(range(len(self.buffer)), k = batch_size, weight=None)
    #print(np.array(self.buffer))
    #mini_batch = np.array(self.buffer)[samples_index]
    mini_batch = [self.buffer[idx] for idx in samples_index]
    importance = self.get_importance(sample_probs[samples_index])

    #mini_batch = random.sample(self.buffer, batch_size)
    states, actions, rewards, next_states = [], [], [], []
    for transition in mini_batch:
        s, a, r, ns = transition
        states.append(s) # s is a tensor not a regular element
        actions.append([a])
        rewards.append([r])
        next_states.append(ns)

    states = torch.stack(states)
    actions = torch.tensor(actions, dtype=torch.float)
    rewards = torch.tensor(rewards, dtype=torch.float)
    next_states = torch.stack(next_states)
    return states, actions, rewards, next_states, torch.tensor(importance), sample_probs

def set_priorities(self, index, error, offset=0.1):
    for i,e in zip(index, error):
        if isinstance(e,np.ndarray):
            e = e[0]
        self.priorities[i] = abs(e) + offset

def __len__(self):
    return len(self.buffer)

```

In [117]:

```

class PERDoubleDQNAgent:
    def __init__(self):
        self.steps = 0
        self.memory = PERMemory(10000)

        # Greedy Epsilon Hyperparams
        self.epsilon = 1.0
        self.epsilon_decay = 0.98
        self.epsilon_min = 0.001

        self.batch_size = 256
        self.adamlr = 0.001
        self.targetlr = 0.01
        self.gamma = 0.98

```

```

self.numOfAction = 20
self.action_array = np.concatenate((( -1 * ( np.exp(np.linspace(0, 1.09, 10)) -
self.policy_net = MLP(3, self.numOfAction, self.adamlr)
self.target_net = MLP(3, self.numOfAction, self.adamlr)
self.target_net.load_state_dict(self.policy_net.state_dict()))

def transform_choice(self, num):
    return self.action_array[int(num)]

def action(self, state):

    sample = np.random.rand()

    if sample > self.epsilon:
        with torch.no_grad():
            choice = float(torch.argmax(self.policy_net(state)).numpy())
            action = self.transform_choice(choice)
    else:
        choice = np.random.choice([n for n in range(self.numOfAction)])
        action = self.transform_choice(choice)

    return choice, action

def optimize(self):
    if len(self.memory) < self.batch_size:
        return

    batch = self.memory.sample(self.batch_size)
    state, action, reward, next_state, importance, sample_index = batch
    action = action.type(torch.int64)

    with torch.no_grad():
        expected_q = self.target_net(next_state).max(1)[0].unsqueeze(1)
        expected = reward + self.gamma * expected_q

    # Actual values - Policy Net
    actual = self.policy_net(state).gather(1,action)
    # Back_propagation
    loss = F.smooth_l1_loss(actual, expected, reduction='none')
    self.memory.set_priorities(sample_index, loss.detach().numpy())
    weighted_loss = torch.multiply(loss, importance)
    self.policy_net.optimizer.zero_grad()
    weighted_loss.mean().backward()
    nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
    self.policy_net.optimizer.step()

    for target_param, param in zip(self.target_net.parameters(), self.policy_net.parameters()):
        target_param.data.copy_(target_param.data * (1.0 - self.targetlr) + param.data * self.policy_net.optimizer.state[param]['weight'])

```

In [118]:

```

perAgent, perdqn_score = train_model_eps(PERDoubleDQNAgent, 300)
plot_tuples([(ddqn_score, "base"), (perdqn_score, "action")], "Compare Dueling DQN")

```

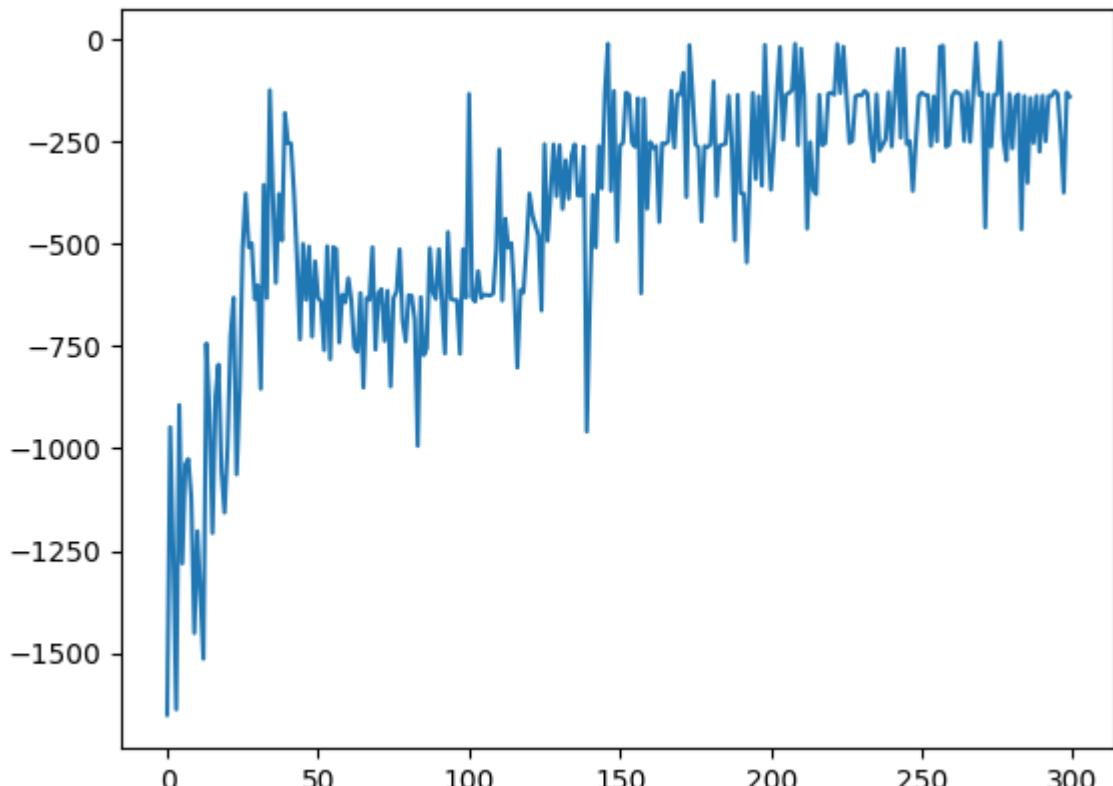
Episode: 0, Score: -1651.2039652587346, Threshold: 1.0  
Episode: 1, Score: -948.1035770469136, Threshold: 0.98  
Episode: 2, Score: -1260.704349962935, Threshold: 0.9603999999999999  
Episode: 3, Score: -1637.9721119972085, Threshold: 0.9411919999999999  
Episode: 4, Score: -893.8395627191157, Threshold: 0.9223681599999999  
Episode: 5, Score: -1281.313962033771, Threshold: 0.9039207967999998  
Episode: 6, Score: -1039.0975574629774, Threshold: 0.8858423808639998  
Episode: 7, Score: -1025.728513747337, Threshold: 0.8681255332467198  
Episode: 8, Score: -1122.5009474180606, Threshold: 0.8507630225817854  
Episode: 9, Score: -1451.4731011313293, Threshold: 0.8337477621301497  
Episode: 10, Score: -1202.2686778507025, Threshold: 0.8170728068875467  
Episode: 11, Score: -1361.2420894156544, Threshold: 0.8007313507497957  
Episode: 12, Score: -1514.2830224281704, Threshold: 0.7847167237347998  
Episode: 13, Score: -743.3064216664995, Threshold: 0.7690223892601038  
Episode: 14, Score: -907.6732117073949, Threshold: 0.7536419414749017  
Episode: 15, Score: -1207.0002258015306, Threshold: 0.7385691026454037  
Episode: 16, Score: -872.4187958993906, Threshold: 0.7237977205924956  
Episode: 17, Score: -795.0602111032525, Threshold: 0.7093217661806457  
Episode: 18, Score: -1057.1136968006458, Threshold: 0.6951353308570327  
Episode: 19, Score: -1156.8274347240067, Threshold: 0.6812326242398921  
Episode: 20, Score: -996.2460367296262, Threshold: 0.6676079717550942  
Episode: 21, Score: -723.9003083829526, Threshold: 0.6542558123199923  
Episode: 22, Score: -631.4678991452868, Threshold: 0.6411706960735924  
Episode: 23, Score: -1063.417378104166, Threshold: 0.6283472821521205  
Episode: 24, Score: -858.0113883632208, Threshold: 0.6157803365090782  
Episode: 25, Score: -507.82145882791946, Threshold: 0.6034647297788965  
Episode: 26, Score: -377.08800291155075, Threshold: 0.5913954351833186  
Episode: 27, Score: -508.138378257455, Threshold: 0.5795675264796523  
Episode: 28, Score: -498.54808850752136, Threshold: 0.5679761759500592  
Episode: 29, Score: -634.7476643737946, Threshold: 0.5566166524310581  
Episode: 30, Score: -601.4691713563008, Threshold: 0.5454843193824369  
Episode: 31, Score: -854.2689751401065, Threshold: 0.5345746329947881  
Episode: 32, Score: -355.67809720719583, Threshold: 0.5238831403348924  
Episode: 33, Score: -632.6162589710071, Threshold: 0.5134054775281945  
Episode: 34, Score: -124.27194774898175, Threshold: 0.5031373679776306  
Episode: 35, Score: -379.8726810662942, Threshold: 0.493074620618078  
Episode: 36, Score: -595.008892774311, Threshold: 0.48321312820571644  
Episode: 37, Score: -378.1633751446076, Threshold: 0.4735488656416021  
Episode: 38, Score: -491.82669433475286, Threshold: 0.46407788832877006  
Episode: 39, Score: -180.09990592193705, Threshold: 0.45479633056219465  
Episode: 40, Score: -254.32662406024895, Threshold: 0.44570040395095073  
Episode: 41, Score: -254.1623724388898, Threshold: 0.4367863958719317  
Episode: 42, Score: -374.70487394354075, Threshold: 0.42805066795449304  
Episode: 43, Score: -535.7709536868442, Threshold: 0.41948965459540316  
Episode: 44, Score: -733.3690436306122, Threshold: 0.4110998615034951  
Episode: 45, Score: -499.9226391681356, Threshold: 0.4028778642734252  
Episode: 46, Score: -636.817241504939, Threshold: 0.39482030698795667  
Episode: 47, Score: -506.3565623538722, Threshold: 0.38692390084819756  
Episode: 48, Score: -726.6624811508069, Threshold: 0.3791854228312336  
Episode: 49, Score: -542.6970346693425, Threshold: 0.37160171437460887  
Episode: 50, Score: -631.9461230771666, Threshold: 0.3641696800871167  
Episode: 51, Score: -640.9943920055316, Threshold: 0.35688628648537435  
Episode: 52, Score: -760.024115433634, Threshold: 0.34974856075566685  
Episode: 53, Score: -505.7437828866954, Threshold: 0.3427535895405535  
Episode: 54, Score: -781.8668009156013, Threshold: 0.33589851774974244  
Episode: 55, Score: -507.7903900166162, Threshold: 0.3291805473947476  
Episode: 56, Score: -512.7986371589154, Threshold: 0.32259693644685267  
Episode: 57, Score: -741.1393300814776, Threshold: 0.3161449977179156  
Episode: 58, Score: -624.1790672339179, Threshold: 0.3098220977635573  
Episode: 59, Score: -643.1702668653063, Threshold: 0.30362565580828615  
Episode: 60, Score: -583.7907517930529, Threshold: 0.2975531426921204  
Episode: 61, Score: -637.4721018216547, Threshold: 0.291602079838278  
Episode: 62, Score: -752.8142750026332, Threshold: 0.2857700382415124  
Episode: 63, Score: -763.9233721975564, Threshold: 0.2800546374766822  
Episode: 64, Score: -620.1455357950533, Threshold: 0.27445354472714856  
Episode: 65, Score: -851.2743148748132, Threshold: 0.2689644738326056

Episode: 66, Score: -632.3805021008045, Threshold: 0.26358518435595346  
Episode: 67, Score: -636.9822444113702, Threshold: 0.25831348066883436  
Episode: 68, Score: -508.2728920714947, Threshold: 0.2531472110554577  
Episode: 69, Score: -758.5106229992614, Threshold: 0.24808426683434853  
Episode: 70, Score: -620.0465002995331, Threshold: 0.24312258149766156  
Episode: 71, Score: -610.4240659847643, Threshold: 0.2382601298677083  
Episode: 72, Score: -737.428047961483, Threshold: 0.23349492727035415  
Episode: 73, Score: -614.6577915859874, Threshold: 0.22882502872494706  
Episode: 74, Score: -848.591481646308, Threshold: 0.22424852815044813  
Episode: 75, Score: -632.459494031948, Threshold: 0.21976355758743915  
Episode: 76, Score: -616.1982798446727, Threshold: 0.21536828643569036  
Episode: 77, Score: -513.3016074070309, Threshold: 0.21106092070697655  
Episode: 78, Score: -687.6359695055631, Threshold: 0.20683970229283702  
Episode: 79, Score: -738.2216261267994, Threshold: 0.20270290824698028  
Episode: 80, Score: -625.5675383883753, Threshold: 0.19864885008204067  
Episode: 81, Score: -626.9652629130193, Threshold: 0.19467587308039985  
Episode: 82, Score: -685.2455013876714, Threshold: 0.19078235561879187  
Episode: 83, Score: -993.6416499187465, Threshold: 0.18696670850641603  
Episode: 84, Score: -630.2395516504916, Threshold: 0.18322737433628772  
Episode: 85, Score: -771.9480597557462, Threshold: 0.17956282684956196  
Episode: 86, Score: -754.569349908279, Threshold: 0.1759715703125707  
Episode: 87, Score: -510.49732181668435, Threshold: 0.1724521389063193  
Episode: 88, Score: -618.1943101290498, Threshold: 0.16900309612819292  
Episode: 89, Score: -634.5502574366695, Threshold: 0.16562303420562907  
Episode: 90, Score: -513.1237215636258, Threshold: 0.16231057352151648  
Episode: 91, Score: -640.319402585872, Threshold: 0.15906436205108615  
Episode: 92, Score: -767.847076072916, Threshold: 0.15588307481006441  
Episode: 93, Score: -471.1029726443966, Threshold: 0.15276541331386312  
Episode: 94, Score: -634.1846058877203, Threshold: 0.14971010504758586  
Episode: 95, Score: -637.410633048889, Threshold: 0.14671590294663414  
Episode: 96, Score: -637.2546512289356, Threshold: 0.14378158488770146  
Episode: 97, Score: -768.8696574283591, Threshold: 0.14090595318994742  
Episode: 98, Score: -513.4697083921866, Threshold: 0.13808783412614847  
Episode: 99, Score: -631.4915934041031, Threshold: 0.13532607744362551  
Episode: 100, Score: -133.36381083063372, Threshold: 0.132619555894753  
Episode: 101, Score: -630.4736723301831, Threshold: 0.12996716477685794  
Episode: 102, Score: -641.336532921488, Threshold: 0.12736782148132078  
Episode: 103, Score: -566.7030360500221, Threshold: 0.12482046505169436  
Episode: 104, Score: -631.9716324048563, Threshold: 0.12232405575066048  
Episode: 105, Score: -624.0784545684444, Threshold: 0.11987757463564727  
Episode: 106, Score: -624.9988001498454, Threshold: 0.11748002314293432  
Episode: 107, Score: -626.3321619327702, Threshold: 0.11513042268007563  
Episode: 108, Score: -621.229101714428, Threshold: 0.11282781422647412  
Episode: 109, Score: -513.927297822861, Threshold: 0.11057125794194463  
Episode: 110, Score: -269.190687792001, Threshold: 0.10835983278310574  
Episode: 111, Score: -638.6929348780017, Threshold: 0.10619263612744362  
Episode: 112, Score: -438.5782741793881, Threshold: 0.10406878340489474  
Episode: 113, Score: -509.4976793465267, Threshold: 0.10198740773679685  
Episode: 114, Score: -497.9995479577645, Threshold: 0.09994765958206091  
Episode: 115, Score: -620.1214131434305, Threshold: 0.0979487063904197  
Episode: 116, Score: -802.5058409264766, Threshold: 0.0959897322626113  
Episode: 117, Score: -612.83487831401, Threshold: 0.09406993761735907  
Episode: 118, Score: -618.920629166785, Threshold: 0.0921885388650119  
Episode: 119, Score: -509.51081703436347, Threshold: 0.09034476808771166  
Episode: 120, Score: -377.00919828977436, Threshold: 0.08853787272595742  
Episode: 121, Score: -429.86985840142944, Threshold: 0.08676711527143827  
Episode: 122, Score: -456.16751301987887, Threshold: 0.0850317729660095  
Episode: 123, Score: -479.08365944473275, Threshold: 0.08333113750668932  
Episode: 124, Score: -662.7393389342042, Threshold: 0.08166451475655553  
Episode: 125, Score: -256.38666643777213, Threshold: 0.08003122446142442  
Episode: 126, Score: -492.7055024410627, Threshold: 0.07843059997219594  
Episode: 127, Score: -379.8416195921399, Threshold: 0.07686198797275202  
Episode: 128, Score: -256.7117186721256, Threshold: 0.07532474821329699  
Episode: 129, Score: -383.5200484222157, Threshold: 0.07381825324903105  
Episode: 130, Score: -258.914988539808, Threshold: 0.07234188818405042  
Episode: 131, Score: -414.4983940913192, Threshold: 0.07089505042036941

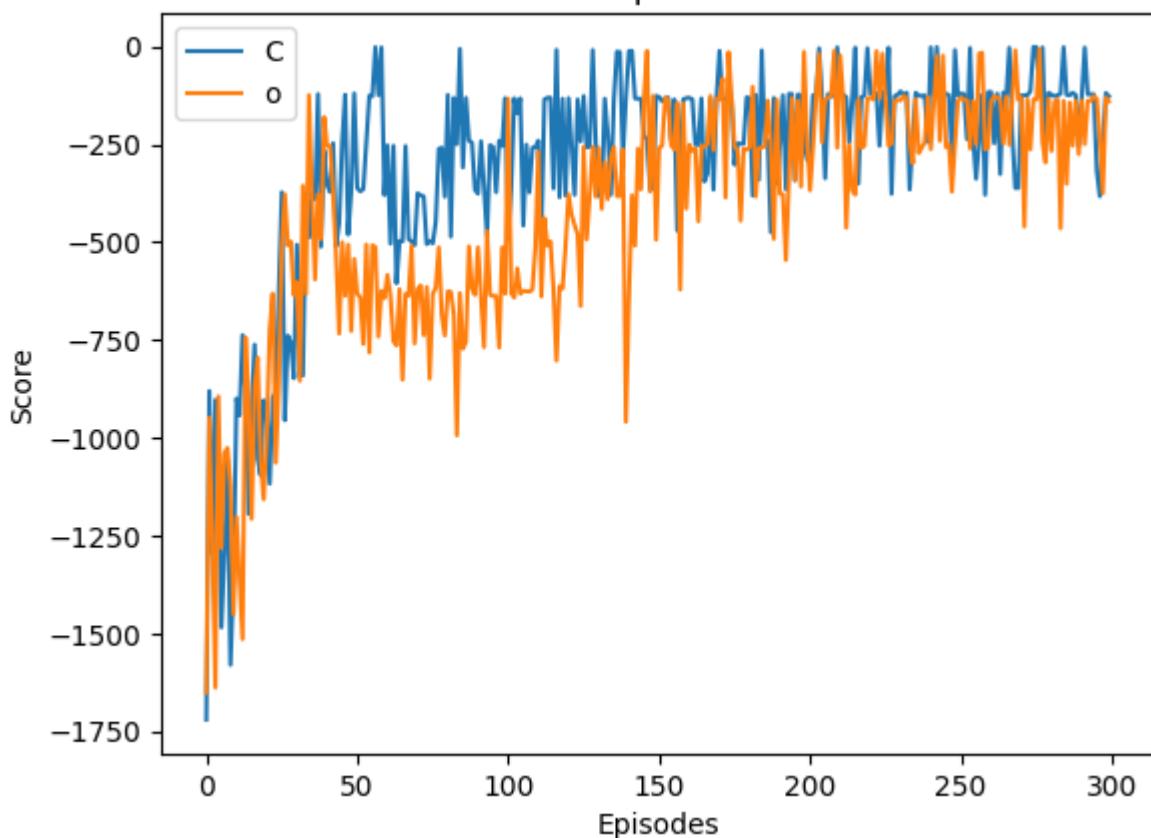
Episode: 132, Score: -295.5446591509165, Threshold: 0.06947714941196202  
Episode: 133, Score: -389.82886335789965, Threshold: 0.06808760642372277  
Episode: 134, Score: -283.2917431962311, Threshold: 0.06672585429524831  
Episode: 135, Score: -256.69332891181693, Threshold: 0.06539133720934334  
Episode: 136, Score: -381.65491567587316, Threshold: 0.06408351046515648  
Episode: 137, Score: -382.2942227681523, Threshold: 0.06280184025585335  
Episode: 138, Score: -262.6086255714217, Threshold: 0.061545803450736285  
Episode: 139, Score: -958.8010112923453, Threshold: 0.060314887381721555  
Episode: 140, Score: -644.5189453265835, Threshold: 0.05910858963408712  
Episode: 141, Score: -380.31808966945357, Threshold: 0.05792641784140538  
Episode: 142, Score: -508.48192657512294, Threshold: 0.05676788948457727  
Episode: 143, Score: -260.8539571098071, Threshold: 0.055632531694885724  
Episode: 144, Score: -364.552253225706, Threshold: 0.054519881060988006  
Episode: 145, Score: -130.29079323445754, Threshold: 0.05342948343976824  
Episode: 146, Score: -10.928981075722255, Threshold: 0.052360893770972874  
Episode: 147, Score: -370.3956972346696, Threshold: 0.05131367589555342  
Episode: 148, Score: -126.2011010035072, Threshold: 0.05028740237764235  
Episode: 149, Score: -493.37599175254593, Threshold: 0.0492816543300895  
Episode: 150, Score: -259.83254394189595, Threshold: 0.04829602124348771  
Episode: 151, Score: -253.09764984050756, Threshold: 0.04733010081861796  
Episode: 152, Score: -130.45347717305515, Threshold: 0.0463834988022456  
Episode: 153, Score: -135.92204478325817, Threshold: 0.04545582882620069  
Episode: 154, Score: -252.57879279282645, Threshold: 0.04454671224967667  
Episode: 155, Score: -262.5577262047303, Threshold: 0.043655778004683135  
Episode: 156, Score: -144.27009486284788, Threshold: 0.04278266244458947  
Episode: 157, Score: -620.814526941771, Threshold: 0.04192700919569768  
Episode: 158, Score: -145.10713560868612, Threshold: 0.04108846901178373  
Episode: 159, Score: -414.07740802069543, Threshold: 0.04026669963154805  
Episode: 160, Score: -250.81230383377894, Threshold: 0.03946136563891709  
Episode: 161, Score: -266.19326672471345, Threshold: 0.03867213832613875  
Episode: 162, Score: -260.92871323929865, Threshold: 0.037898695559615975  
Episode: 163, Score: -446.5176712962544, Threshold: 0.03714072164842366  
Episode: 164, Score: -253.93364171869752, Threshold: 0.03639790721545518  
Episode: 165, Score: -255.6203641700711, Threshold: 0.035669949071146075  
Episode: 166, Score: -249.5572551252534, Threshold: 0.034956550089723155  
Episode: 167, Score: -126.05114494019831, Threshold: 0.03425741908792869  
Episode: 168, Score: -264.6244052652257, Threshold: 0.03357227070617012  
Episode: 169, Score: -133.73304887687314, Threshold: 0.032900825292046715  
Episode: 170, Score: -134.4174551736791, Threshold: 0.03224280878620578  
Episode: 171, Score: -81.90515712543856, Threshold: 0.031597952610481664  
Episode: 172, Score: -385.5846209581132, Threshold: 0.03096599355827203  
Episode: 173, Score: -14.266714428062839, Threshold: 0.030346673687106588  
Episode: 174, Score: -130.40912178914957, Threshold: 0.029739740213364455  
Episode: 175, Score: -256.64377483653567, Threshold: 0.029144945409097164  
Episode: 176, Score: -264.1431808961133, Threshold: 0.02856204650091522  
Episode: 177, Score: -444.8516135438406, Threshold: 0.027990805570896914  
Episode: 178, Score: -260.9810919870956, Threshold: 0.027430989459478975  
Episode: 179, Score: -263.458711373119, Threshold: 0.026882369670289397  
Episode: 180, Score: -257.0217990085019, Threshold: 0.026344722276883607  
Episode: 181, Score: -102.85082595587099, Threshold: 0.025817827831345935  
Episode: 182, Score: -383.0776390210768, Threshold: 0.025301471274719014  
Episode: 183, Score: -260.36907408564315, Threshold: 0.024795441849224633  
Episode: 184, Score: -257.50052755655065, Threshold: 0.02429953301224014  
Episode: 185, Score: -255.02398015696588, Threshold: 0.02381354235199534  
Episode: 186, Score: -137.88285442808098, Threshold: 0.02333727150495543  
Episode: 187, Score: -251.53333408982482, Threshold: 0.02287052607485632  
Episode: 188, Score: -491.44445088374437, Threshold: 0.022413115553359196  
Episode: 189, Score: -135.7009327044386, Threshold: 0.021964853242292012  
Episode: 190, Score: -376.0802752941553, Threshold: 0.02152555617744617  
Episode: 191, Score: -377.5116291187134, Threshold: 0.021095045053897248  
Episode: 192, Score: -545.412797200951, Threshold: 0.020673144152819303  
Episode: 193, Score: -336.2517847020116, Threshold: 0.02025968126976292  
Episode: 194, Score: -131.14509583723265, Threshold: 0.01985448764436766  
Episode: 195, Score: -341.4316247607168, Threshold: 0.019457397891480308  
Episode: 196, Score: -139.12002995907667, Threshold: 0.0190682499336507  
Episode: 197, Score: -357.81478835835094, Threshold: 0.018686884934977686

Episode: 198, Score: -13.58230029610016, Threshold: 0.01831314723627813  
Episode: 199, Score: -235.78494105895606, Threshold: 0.017946884291552567  
Episode: 200, Score: -367.468000217865, Threshold: 0.017587946605721515  
Episode: 201, Score: -252.96050140303998, Threshold: 0.017236187673607083  
Episode: 202, Score: -125.31303590754611, Threshold: 0.016891463920134943  
Episode: 203, Score: -18.70181357972518, Threshold: 0.016553634641732245  
Episode: 204, Score: -245.1189597126025, Threshold: 0.0162225619488976  
Episode: 205, Score: -135.4834771136489, Threshold: 0.015898110709919647  
Episode: 206, Score: -131.14597845557716, Threshold: 0.015580148495721254  
Episode: 207, Score: -124.12006107745817, Threshold: 0.015268545525806828  
Episode: 208, Score: -11.083254118350286, Threshold: 0.014963174615290691  
Episode: 209, Score: -258.5543102099279, Threshold: 0.014663911122984877  
Episode: 210, Score: -22.701567791222416, Threshold: 0.01437063290052518  
Episode: 211, Score: -136.04162210356668, Threshold: 0.014083220242514675  
Episode: 212, Score: -462.8616377850184, Threshold: 0.013801555837664382  
Episode: 213, Score: -251.25639476875114, Threshold: 0.013525524720911095  
Episode: 214, Score: -367.77323353978437, Threshold: 0.013255014226492874  
Episode: 215, Score: -378.370065921107, Threshold: 0.012989913941963017  
Episode: 216, Score: -135.50782646676606, Threshold: 0.012730115663123756  
Episode: 217, Score: -258.88251397015915, Threshold: 0.012475513349861281  
Episode: 218, Score: -253.953891355764, Threshold: 0.012226003082864055  
Episode: 219, Score: -132.96524669724502, Threshold: 0.011981483021206774  
Episode: 220, Score: -130.59568527056967, Threshold: 0.011741853360782638  
Episode: 221, Score: -135.32571607377562, Threshold: 0.011507016293566985  
Episode: 222, Score: -11.545555096792466, Threshold: 0.011276875967695645  
Episode: 223, Score: -131.2870814600293, Threshold: 0.011051338448341731  
Episode: 224, Score: -18.01098375161623, Threshold: 0.010830311679374895  
Episode: 225, Score: -134.24460849161719, Threshold: 0.010613705445787398  
Episode: 226, Score: -252.9679581888489, Threshold: 0.01040143133687165  
Episode: 227, Score: -248.14877893989527, Threshold: 0.010193402710134216  
Episode: 228, Score: -139.39291902886373, Threshold: 0.009989534655931531  
Episode: 229, Score: -136.12050492774569, Threshold: 0.0097897439628129  
Episode: 230, Score: -137.4893513499845, Threshold: 0.009593949083556642  
Episode: 231, Score: -125.6619982283113, Threshold: 0.009402070101885509  
Episode: 232, Score: -133.3530989447405, Threshold: 0.009214028699847799  
Episode: 233, Score: -243.98825420781614, Threshold: 0.009029748125850843  
Episode: 234, Score: -298.2970874712505, Threshold: 0.008849153163333826  
Episode: 235, Score: -134.1795167998195, Threshold: 0.008672170100067149  
Episode: 236, Score: -272.00767533965154, Threshold: 0.008498726698065806  
Episode: 237, Score: -258.8824696043236, Threshold: 0.00832875216410449  
Episode: 238, Score: -243.47402631695394, Threshold: 0.0081621771208224  
Episode: 239, Score: -128.89650915726403, Threshold: 0.007998933578405951  
Episode: 240, Score: -261.9868416717368, Threshold: 0.007838954906837833  
Episode: 241, Score: -142.72370402675818, Threshold: 0.007682175808701076  
Episode: 242, Score: -23.39570131539674, Threshold: 0.007528532292527055  
Episode: 243, Score: -240.60406397659702, Threshold: 0.007377961646676513  
Episode: 244, Score: -23.17579401105966, Threshold: 0.007230402413742983  
Episode: 245, Score: -256.48286069364303, Threshold: 0.007085794365468123  
Episode: 246, Score: -249.23018348695138, Threshold: 0.00694407847815876  
Episode: 247, Score: -370.55049262428025, Threshold: 0.0068051969085955845  
Episode: 248, Score: -249.66779649119437, Threshold: 0.006669092970423673  
Episode: 249, Score: -138.1379109360947, Threshold: 0.00653571111015199  
Episode: 250, Score: -130.21009201851237, Threshold: 0.006404996888794895  
Episode: 251, Score: -136.09730620062095, Threshold: 0.006276896951018997  
Episode: 252, Score: -136.83994589334094, Threshold: 0.006151359011998617  
Episode: 253, Score: -260.689321871402, Threshold: 0.006028331831758644  
Episode: 254, Score: -139.23910119245812, Threshold: 0.005907765195123471  
Episode: 255, Score: -249.20275447153372, Threshold: 0.005789609891221002  
Episode: 256, Score: -18.164755922263158, Threshold: 0.005673817693396582  
Episode: 257, Score: -15.727826561407264, Threshold: 0.00556034133952865  
Episode: 258, Score: -262.7018900312227, Threshold: 0.005449134512738076  
Episode: 259, Score: -257.51682573934926, Threshold: 0.0053401518224833145  
Episode: 260, Score: -139.66545807762438, Threshold: 0.005233348786033648  
Episode: 261, Score: -126.8918512682359, Threshold: 0.005128681810312975  
Episode: 262, Score: -131.24264972964633, Threshold: 0.005026108174106716  
Episode: 263, Score: -135.255923297573, Threshold: 0.004925586010624581

Episode: 264, Score: -248.31794660209576, Threshold: 0.004827074290412089  
Episode: 265, Score: -127.54198634989251, Threshold: 0.004730532804603847  
Episode: 266, Score: -250.70040675412062, Threshold: 0.00463592214851177  
Episode: 267, Score: -134.94023738209572, Threshold: 0.004543203705541535  
Episode: 268, Score: -9.426577904041189, Threshold: 0.004452339631430704  
Episode: 269, Score: -135.90742262557603, Threshold: 0.00436329283880209  
Episode: 270, Score: -130.44965697594264, Threshold: 0.004276026982026048  
Episode: 271, Score: -459.8165549494425, Threshold: 0.004190506442385527  
Episode: 272, Score: -134.52726380081435, Threshold: 0.004106696313537816  
Episode: 273, Score: -262.3296543427466, Threshold: 0.00402456238726706  
Episode: 274, Score: -134.99377641456675, Threshold: 0.003944071139521719  
Episode: 275, Score: -134.99465888851356, Threshold: 0.0038651897167312848  
Episode: 276, Score: -6.585724519813376, Threshold: 0.003787885922396659  
Episode: 277, Score: -243.72244913672517, Threshold: 0.003712128203948726  
Episode: 278, Score: -296.11305010596834, Threshold: 0.0036378856398697514  
Episode: 279, Score: -133.46076256755106, Threshold: 0.0035651279270723563  
Episode: 280, Score: -266.586430292509, Threshold: 0.0034938253685309092  
Episode: 281, Score: -141.20675826425722, Threshold: 0.003423948861160291  
Episode: 282, Score: -134.60377076362445, Threshold: 0.0033554698839370852  
Episode: 283, Score: -464.265988301513, Threshold: 0.0032883604862583436  
Episode: 284, Score: -138.72096416652334, Threshold: 0.0032225932765331767  
Episode: 285, Score: -350.6469551669259, Threshold: 0.003158141411002513  
Episode: 286, Score: -143.20503835014568, Threshold: 0.003094978582782463  
Episode: 287, Score: -253.30688121472545, Threshold: 0.0030330790111268136  
Episode: 288, Score: -138.353471122459, Threshold: 0.0029724174309042773  
Episode: 289, Score: -274.69148944511795, Threshold: 0.002912969082286192  
Episode: 290, Score: -137.93262766915538, Threshold: 0.002854709700640468  
Episode: 291, Score: -249.23320047055472, Threshold: 0.0027976155066276585  
Episode: 292, Score: -138.5677637650853, Threshold: 0.0027416631964951054  
Episode: 293, Score: -138.42799750771007, Threshold: 0.0026868299325652034  
Episode: 294, Score: -126.4154165919666, Threshold: 0.002633093333913899  
Episode: 295, Score: -133.40734882975187, Threshold: 0.002580431467235621  
Episode: 296, Score: -234.57599923141106, Threshold: 0.0025288228378909086  
Episode: 297, Score: -374.6446300881245, Threshold: 0.0024782463811330903  
Episode: 298, Score: -129.87765637389398, Threshold: 0.0024286814535104282  
Episode: 299, Score: -140.18225661236002, Threshold: 0.0023801078244402194

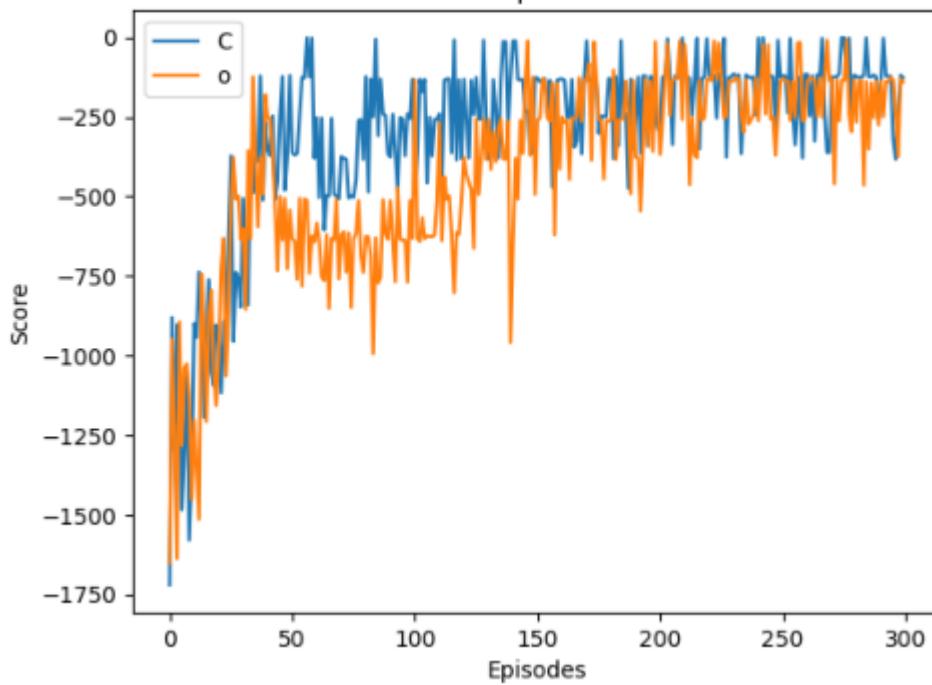


Comparison



```
In [122]: torch.save(perAgent.policy_net.state_dict(), "./rl_models/per_dqn.pt")
showImage("./Img/RL/per_ddqn.png")
```

Comparison



Observations:

### Training Time

- Prioritised Replay is significant slower in terms of training as it add more model complexity
- The time take for 300 episodes was around 1.5 times longer

### Performance

- As seen on the graph PER did not provide any groundbreaking results

- It took longer to converge
- At the end had similar performance to a Double DQN

## Limitations of PER in the Pendulum Problem

In the case of the Pendulum problem, certain characteristics of the environment might limit the effectiveness of Prioritized Replay:

### 1. Lack of Sparse Rewards

The Pendulum environment typically provides continuous and non-sparse rewards. Since Prioritized Replay is particularly beneficial in problems with sparse and informative rewards, its application in the Pendulum problem may not yield significant improvements.

### 2. Continuous Action Space

The Pendulum problem involves a continuous action space, making it inherently different from discrete action space problems where prioritized replay has shown more impact. The continuous nature of actions might reduce the influence of prioritization on learning.

### 3. Lack of Temporal Dependencies

In some environments, prioritized replay helps to address temporal dependencies in the data, ensuring that crucial experiences are replayed and learned more frequently. In the Pendulum problem, where the state transitions are less dependent on past experiences, prioritized replay may not provide a substantial advantage.

### 4. Lack of Environmental Complexity

The pendulum environment is relatively simple compared to others like Atari games or snakes and ladders. In those environments, sampling the right info is more important and thus, PER might be more successful.

## Noisy Networks

Noisy Networks introduce stochasticity directly into the parameters of a neural network, allowing the model to explore action spaces more effectively.

**Key Benefits:**

### 1. Exploration Enhancement:

- Traditional exploration strategies, such as epsilon-greedy, might not be sufficient in complex environments. Noisy Networks introduce random perturbations into the network's weights, allowing the agent to explore a broader range of actions.

### 2. Parametric Noise:

- Noisy Layers model exploration as parameterized noise added to the activations of the layer. This noise is adaptive and trainable, providing the network with the flexibility to learn the level of exploration required for optimal performance.

### 3. Reducing Overestimation Bias:

- Noisy Networks can help mitigate overestimation bias, a common challenge in DDQN. By introducing stochasticity during action selection, the agent becomes less reliant on deterministic Q-values, resulting in a more accurate estimation of state-action values.

**Implementation in DDQN:**

To integrate Noisy Networks into our DDQN implementation, we can just include Noisy Layers to the fully connected layers of both the policy and target networks. This can be achieved using PyTorch's NoisyLinear layers.

In [119...]

```
class DDQN_MLP_Noisy(nn.Module):
    def __init__(self, inp, out, lr):
        super(DDQN_MLP_Noisy, self).__init__()

        self.layer1 = nn.Linear(inp, 32)
        self.layer2 = nn.Linear(32, 64)

        self.value_stream = nn.Sequential(
            nn.Linear(64, 64),
            nn.LeakyReLU(inplace=True),
            NoisyLinear(64, 1)
        )

        self.advantage_stream = nn.Sequential(
            nn.Linear(64, 64),
            nn.LeakyReLU(inplace=True),
            NoisyLinear(64, out)
        )

        self.lr = lr
        self.optimizer = optim.Adam(self.parameters(), lr=self.lr)

    def forward(self, x):
        x = self.layer1(x)
        x = F.leaky_relu(x)
        x = self.layer2(x)
        x = F.leaky_relu(x)

        values = self.value_stream(x)
        advantages = self.advantage_stream(x)
        q_values = values + advantages - advantages.mean()

    return q_values
```

In [120...]

```
class NoisyDoubleDQNAgent:
    def __init__(self):
        self.steps = 0
        self.memory = Memory(10000)

        # Greedy Epsilon Hyperparams
        self.epsilon = 1.0
        self.epsilon_decay = 0.98
        self.epsilon_min = 0.001

        self.batch_size = 256
        self.adamlr = 0.001
        self.targetlr = 0.01
        self.gamma = 0.98

        self.numOfAction = 20
        self.action_array = np.concatenate((-1 * (np.exp(np.linspace(0, 1.09, 10)) - 1),))

        self.policy_net = DDQN_MLP_Noisy(3, self.numOfAction, self.adamlr)
        self.target_net = DDQN_MLP_Noisy(3, self.numOfAction, self.adamlr)
        self.target_net.load_state_dict(self.policy_net.state_dict())

    def transform_choice(self, num):
        return self.action_array[int(num)]
```

```

def action(self,state):
    sample = np.random.rand()

    if sample > self.epsilon:
        with torch.no_grad():
            choice = float(torch.argmax(self.policy_net(state)).numpy())
            action = self.transform_choice(choice)
    else:
        choice = np.random.choice([n for n in range(self.numOfAction)])
        action = self.transform_choice(choice)

    return choice, action

def optimize(self):
    if len(self.memory) < self.batch_size:
        return

    batch = self.memory.sample(self.batch_size)
    state, action, reward, next_state = batch
    action = action.type(torch.int64)

    with torch.no_grad():
        expected_q = self.target_net(next_state).max(1)[0].unsqueeze(1)
        expected = reward + self.gamma * expected_q

    # Actual values - Policy Net
    actual = self.policy_net(state).gather(1,action)

    # Back_propagation
    loss = F.smooth_l1_loss(actual, expected)
    self.policy_net.optimizer.zero_grad()
    loss.mean().backward()
    nn.utils.clip_grad_norm_(self.policy_net.parameters(), 0.5)
    self.policy_net.optimizer.step()

    for target_param,param in zip(self.target_net.parameters(),self.policy_net.pa
        target_param.data.copy_(target_param.data * (1.0 - self.targetlr) + param

```

In [121...]: noisyAgent, noisydqn\_score = train\_model\_eps(NoisyDoubleDQNAgent,300)  
plot\_tuples([(ddqn\_score,"base"),(noisydqn\_score,"noisy")],"Compare Dueling DQN")

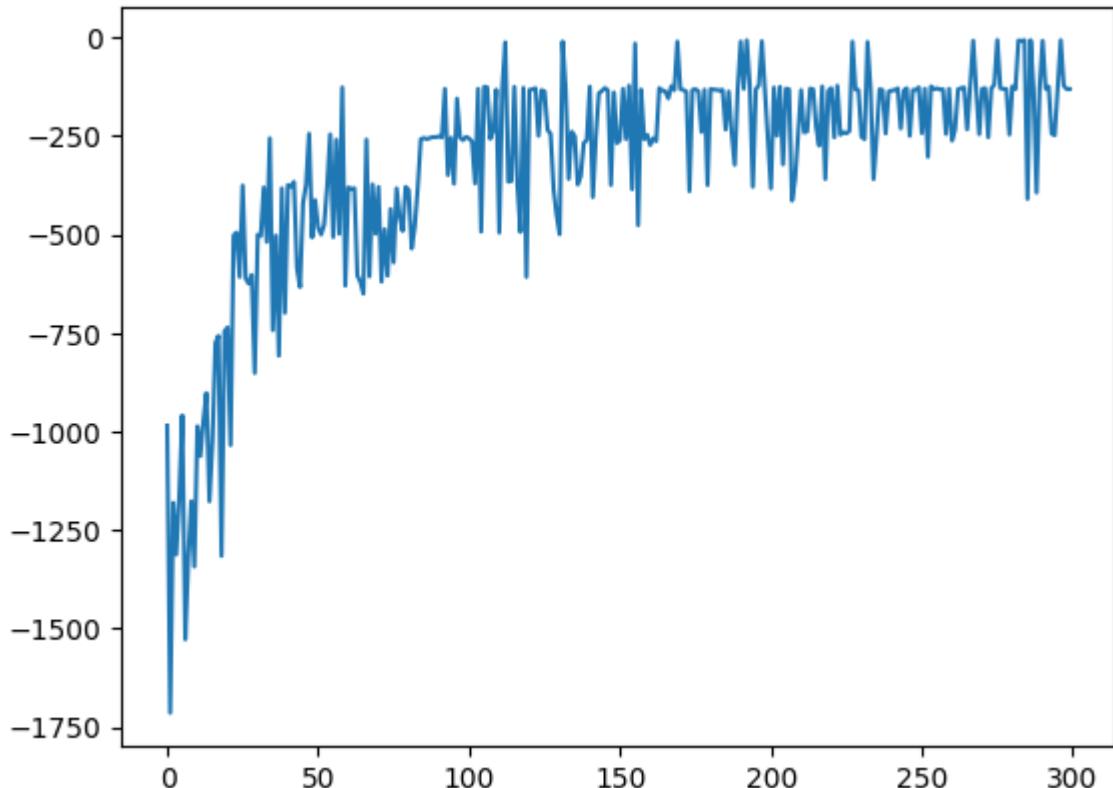
Episode: 0, Score: -985.0381655892998, Threshold: 1.0  
Episode: 1, Score: -1714.2130167318405, Threshold: 0.98  
Episode: 2, Score: -1181.034698308323, Threshold: 0.9603999999999999  
Episode: 3, Score: -1312.6121774962191, Threshold: 0.9411919999999999  
Episode: 4, Score: -1164.8999911553904, Threshold: 0.9223681599999999  
Episode: 5, Score: -959.0617069234175, Threshold: 0.9039207967999998  
Episode: 6, Score: -1527.6393912435121, Threshold: 0.8858423808639998  
Episode: 7, Score: -1320.233344275734, Threshold: 0.8681255332467198  
Episode: 8, Score: -1176.9695636244257, Threshold: 0.8507630225817854  
Episode: 9, Score: -1342.8412274654556, Threshold: 0.8337477621301497  
Episode: 10, Score: -986.960926892078, Threshold: 0.8170728068875467  
Episode: 11, Score: -1062.4448026404207, Threshold: 0.8007313507497957  
Episode: 12, Score: -972.3207580584801, Threshold: 0.7847167237347998  
Episode: 13, Score: -902.5422976890397, Threshold: 0.7690223892601038  
Episode: 14, Score: -1177.3752235010872, Threshold: 0.7536419414749017  
Episode: 15, Score: -1024.3545401697556, Threshold: 0.7385691026454037  
Episode: 16, Score: -774.7204592429432, Threshold: 0.7237977205924956  
Episode: 17, Score: -757.1400016347809, Threshold: 0.7093217661806457  
Episode: 18, Score: -1315.8327585101079, Threshold: 0.6951353308570327  
Episode: 19, Score: -745.6832131073966, Threshold: 0.6812326242398921  
Episode: 20, Score: -735.7144237728293, Threshold: 0.6676079717550942  
Episode: 21, Score: -1034.7173241163061, Threshold: 0.6542558123199923  
Episode: 22, Score: -502.4447797195432, Threshold: 0.6411706960735924  
Episode: 23, Score: -495.0951185756293, Threshold: 0.6283472821521205  
Episode: 24, Score: -607.6075211399731, Threshold: 0.6157803365090782  
Episode: 25, Score: -375.6448036480414, Threshold: 0.6034647297788965  
Episode: 26, Score: -611.3372514124007, Threshold: 0.5913954351833186  
Episode: 27, Score: -624.5773879331938, Threshold: 0.5795675264796523  
Episode: 28, Score: -602.5785525496148, Threshold: 0.5679761759500592  
Episode: 29, Score: -851.4583977552861, Threshold: 0.5566166524310581  
Episode: 30, Score: -500.4302282047774, Threshold: 0.5454843193824369  
Episode: 31, Score: -503.57280895897645, Threshold: 0.5345746329947881  
Episode: 32, Score: -380.48299983942684, Threshold: 0.5238831403348924  
Episode: 33, Score: -518.8670051936493, Threshold: 0.5134054775281945  
Episode: 34, Score: -255.55444685317934, Threshold: 0.5031373679776306  
Episode: 35, Score: -742.9890228785888, Threshold: 0.493074620618078  
Episode: 36, Score: -502.83821203857804, Threshold: 0.48321312820571644  
Episode: 37, Score: -807.6301453750666, Threshold: 0.4735488656416021  
Episode: 38, Score: -382.6656676284146, Threshold: 0.46407788832877006  
Episode: 39, Score: -698.781811298458, Threshold: 0.45479633056219465  
Episode: 40, Score: -374.45819541900386, Threshold: 0.44570040395095073  
Episode: 41, Score: -377.7875073710148, Threshold: 0.4367863958719317  
Episode: 42, Score: -366.4495303692085, Threshold: 0.42805066795449304  
Episode: 43, Score: -585.5634767528544, Threshold: 0.41948965459540316  
Episode: 44, Score: -633.2993429321589, Threshold: 0.4110998615034951  
Episode: 45, Score: -420.26390949203125, Threshold: 0.4028778642734252  
Episode: 46, Score: -374.11065791233887, Threshold: 0.39482030698795667  
Episode: 47, Score: -244.17120346456332, Threshold: 0.38692390084819756  
Episode: 48, Score: -507.58597541596794, Threshold: 0.3791854228312336  
Episode: 49, Score: -412.5521138964321, Threshold: 0.37160171437460887  
Episode: 50, Score: -482.4737035620013, Threshold: 0.3641696800871167  
Episode: 51, Score: -499.98894531350254, Threshold: 0.35688628648537435  
Episode: 52, Score: -474.77829673242735, Threshold: 0.34974856075566685  
Episode: 53, Score: -377.18653676536843, Threshold: 0.3427535895405535  
Episode: 54, Score: -245.71794911719303, Threshold: 0.33589851774974244  
Episode: 55, Score: -506.5332412492111, Threshold: 0.3291805473947476  
Episode: 56, Score: -259.2495378709567, Threshold: 0.32259693644685267  
Episode: 57, Score: -498.24853380086694, Threshold: 0.3161449977179156  
Episode: 58, Score: -126.5769811923891, Threshold: 0.3098220977635573  
Episode: 59, Score: -629.9266023446131, Threshold: 0.30362565580828615  
Episode: 60, Score: -380.6693934742223, Threshold: 0.2975531426921204  
Episode: 61, Score: -384.2517919202563, Threshold: 0.291602079838278  
Episode: 62, Score: -382.41694638659095, Threshold: 0.2857700382415124  
Episode: 63, Score: -604.0355749921, Threshold: 0.2800546374766822  
Episode: 64, Score: -619.9633426471245, Threshold: 0.27445354472714856  
Episode: 65, Score: -650.0238178067085, Threshold: 0.2689644738326056

Episode: 66, Score: -259.05836734146243, Threshold: 0.26358518435595346  
Episode: 67, Score: -606.2470403790345, Threshold: 0.25831348066883436  
Episode: 68, Score: -372.67596730235636, Threshold: 0.2531472110554577  
Episode: 69, Score: -498.2325317437374, Threshold: 0.24808426683434853  
Episode: 70, Score: -379.17543810112795, Threshold: 0.24312258149766156  
Episode: 71, Score: -619.7528608646633, Threshold: 0.2382601298677083  
Episode: 72, Score: -486.59438828419525, Threshold: 0.23349492727035415  
Episode: 73, Score: -604.6383958208008, Threshold: 0.22882502872494706  
Episode: 74, Score: -435.36696080026655, Threshold: 0.22424852815044813  
Episode: 75, Score: -570.9190254116813, Threshold: 0.21976355758743915  
Episode: 76, Score: -382.8643286167003, Threshold: 0.21536828643569036  
Episode: 77, Score: -455.3635717396139, Threshold: 0.21106092070697655  
Episode: 78, Score: -490.96297214014044, Threshold: 0.20683970229283702  
Episode: 79, Score: -379.149568905385, Threshold: 0.20270290824698028  
Episode: 80, Score: -387.08103425687517, Threshold: 0.19864885008204067  
Episode: 81, Score: -535.3816061546571, Threshold: 0.19467587308039985  
Episode: 82, Score: -480.6609420941678, Threshold: 0.19078235561879187  
Episode: 83, Score: -381.89954879569945, Threshold: 0.18696670850641603  
Episode: 84, Score: -258.173985098138, Threshold: 0.18322737433628772  
Episode: 85, Score: -254.95960055587122, Threshold: 0.17956282684956196  
Episode: 86, Score: -257.8237033612945, Threshold: 0.1759715703125707  
Episode: 87, Score: -254.9746019697896, Threshold: 0.1724521389063193  
Episode: 88, Score: -252.81728920620904, Threshold: 0.16900309612819292  
Episode: 89, Score: -252.77543790763912, Threshold: 0.16562303420562907  
Episode: 90, Score: -250.86684702795299, Threshold: 0.16231057352151648  
Episode: 91, Score: -253.52227593034826, Threshold: 0.15906436205108615  
Episode: 92, Score: -130.2004025138604, Threshold: 0.15588307481006441  
Episode: 93, Score: -349.01217136986395, Threshold: 0.15276541331386312  
Episode: 94, Score: -254.202157324977, Threshold: 0.14971010504758586  
Episode: 95, Score: -371.1192717168818, Threshold: 0.14671590294663414  
Episode: 96, Score: -154.97712572062866, Threshold: 0.14378158488770146  
Episode: 97, Score: -253.27139680139794, Threshold: 0.14090595318994742  
Episode: 98, Score: -260.5283169859525, Threshold: 0.13808783412614847  
Episode: 99, Score: -252.14835380000474, Threshold: 0.13532607744362551  
Episode: 100, Score: -255.98914681220276, Threshold: 0.132619555894753  
Episode: 101, Score: -264.2485898751279, Threshold: 0.12996716477685794  
Episode: 102, Score: -370.2769166680042, Threshold: 0.12736782148132078  
Episode: 103, Score: -130.07117150075388, Threshold: 0.12482046505169436  
Episode: 104, Score: -492.69634938805297, Threshold: 0.12232405575066048  
Episode: 105, Score: -124.75327006896637, Threshold: 0.11987757463564727  
Episode: 106, Score: -126.2415189114541, Threshold: 0.11748002314293432  
Episode: 107, Score: -257.1853547358643, Threshold: 0.11513042268007563  
Episode: 108, Score: -235.87402204901463, Threshold: 0.11282781422647412  
Episode: 109, Score: -132.41020451124203, Threshold: 0.11057125794194463  
Episode: 110, Score: -495.19985869745756, Threshold: 0.10835983278310574  
Episode: 111, Score: -130.27262399024022, Threshold: 0.10619263612744362  
Episode: 112, Score: -12.185131967079338, Threshold: 0.10406878340489474  
Episode: 113, Score: -365.79478336102255, Threshold: 0.10198740773679685  
Episode: 114, Score: -364.2891754565407, Threshold: 0.09994765958206091  
Episode: 115, Score: -125.09782704152705, Threshold: 0.0979487063904197  
Episode: 116, Score: -371.19024232125196, Threshold: 0.0959897322626113  
Episode: 117, Score: -493.55615484111075, Threshold: 0.09406993761735907  
Episode: 118, Score: -127.62332306868032, Threshold: 0.0921885388650119  
Episode: 119, Score: -607.6203945247914, Threshold: 0.09034476808771166  
Episode: 120, Score: -132.93808690445132, Threshold: 0.08853787272595742  
Episode: 121, Score: -131.5537025109609, Threshold: 0.08676711527143827  
Episode: 122, Score: -128.6377724212288, Threshold: 0.0850317729660095  
Episode: 123, Score: -249.45707065448588, Threshold: 0.08333113750668932  
Episode: 124, Score: -133.74771931411703, Threshold: 0.08166451475655553  
Episode: 125, Score: -136.7366699639984, Threshold: 0.08003122446142442  
Episode: 126, Score: -234.30568283559128, Threshold: 0.07843059997219594  
Episode: 127, Score: -244.2287239188898, Threshold: 0.07686198797275202  
Episode: 128, Score: -383.811090459301, Threshold: 0.07532474821329699  
Episode: 129, Score: -450.19406000847425, Threshold: 0.07381825324903105  
Episode: 130, Score: -499.2416720644821, Threshold: 0.07234188818405042  
Episode: 131, Score: -10.25156869795649, Threshold: 0.07089505042036941

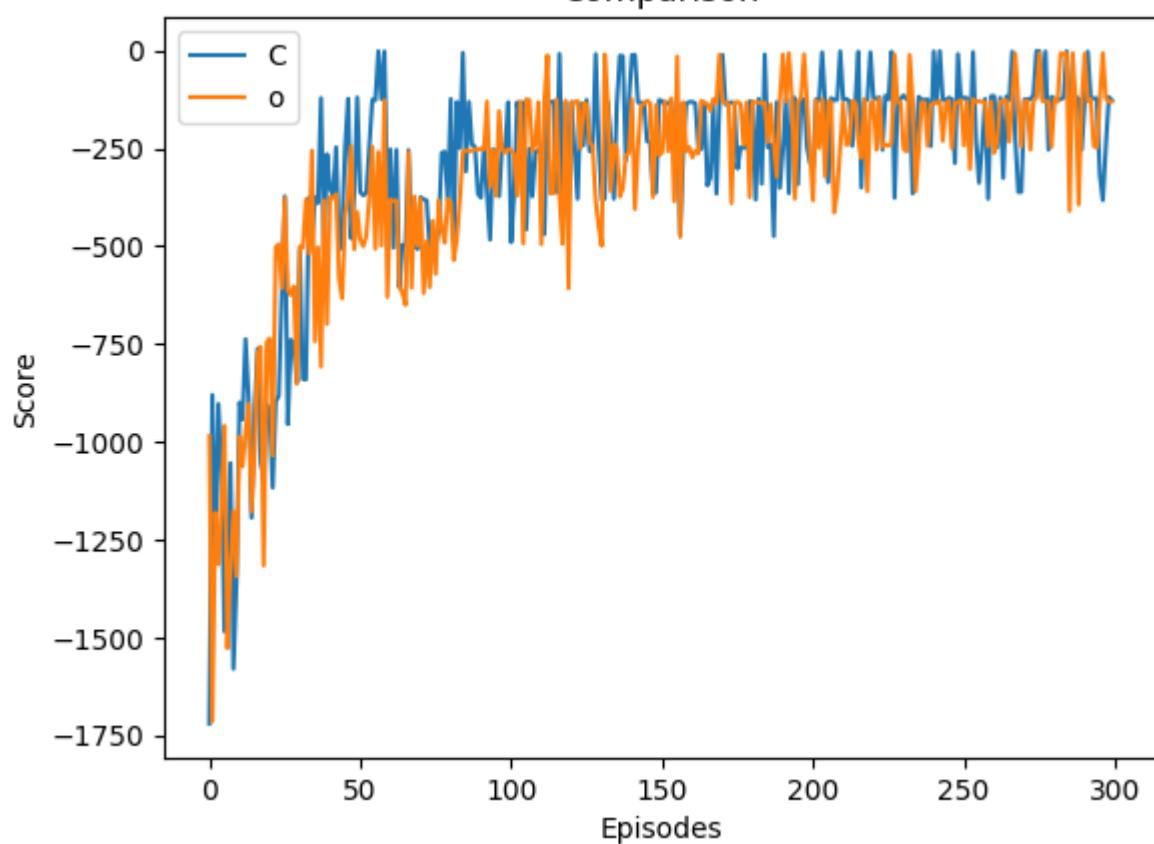
Episode: 132, Score: -132.40613395429082, Threshold: 0.06947714941196202  
Episode: 133, Score: -358.9323381905113, Threshold: 0.06808760642372277  
Episode: 134, Score: -238.97845561394627, Threshold: 0.06672585429524831  
Episode: 135, Score: -250.52516565980045, Threshold: 0.06539133720934334  
Episode: 136, Score: -373.21714391784593, Threshold: 0.06408351046515648  
Episode: 137, Score: -351.7304780795335, Threshold: 0.06280184025585335  
Episode: 138, Score: -267.9230941176717, Threshold: 0.061545803450736285  
Episode: 139, Score: -260.1068369822401, Threshold: 0.060314887381721555  
Episode: 140, Score: -124.4539301973621, Threshold: 0.05910858963408712  
Episode: 141, Score: -405.01339270770546, Threshold: 0.05792641784140538  
Episode: 142, Score: -250.76821935873832, Threshold: 0.05676788948457727  
Episode: 143, Score: -141.4512228549122, Threshold: 0.055632531694885724  
Episode: 144, Score: -135.61173266179534, Threshold: 0.054519881060988006  
Episode: 145, Score: -128.10270879052578, Threshold: 0.05342948343976824  
Episode: 146, Score: -134.11203051964475, Threshold: 0.052360893770972874  
Episode: 147, Score: -374.31996670395773, Threshold: 0.05131367589555342  
Episode: 148, Score: -139.30248746277928, Threshold: 0.05028740237764235  
Episode: 149, Score: -267.79503178042376, Threshold: 0.0492816543300895  
Episode: 150, Score: -260.3826566890037, Threshold: 0.04829602124348771  
Episode: 151, Score: -130.20085312725368, Threshold: 0.04733010081861796  
Episode: 152, Score: -257.56650433169716, Threshold: 0.0463834988022456  
Episode: 153, Score: -122.12267417388199, Threshold: 0.04545582882620069  
Episode: 154, Score: -384.99271568601665, Threshold: 0.04454671224967667  
Episode: 155, Score: -15.293908817466384, Threshold: 0.043655778004683135  
Episode: 156, Score: -476.3325700685102, Threshold: 0.04278266244458947  
Episode: 157, Score: -132.15045103140832, Threshold: 0.04192700919569768  
Episode: 158, Score: -257.74403597017545, Threshold: 0.04108846901178373  
Episode: 159, Score: -246.92641641946565, Threshold: 0.04026669963154805  
Episode: 160, Score: -272.8741510587584, Threshold: 0.03946136563891709  
Episode: 161, Score: -255.82913421939472, Threshold: 0.03867213832613875  
Episode: 162, Score: -262.9282394342292, Threshold: 0.037898695559615975  
Episode: 163, Score: -128.3235122580448, Threshold: 0.03714072164842366  
Episode: 164, Score: -134.2143518560554, Threshold: 0.03639790721545518  
Episode: 165, Score: -136.54814522967135, Threshold: 0.035669949071146075  
Episode: 166, Score: -154.29972988959398, Threshold: 0.034956550089723155  
Episode: 167, Score: -123.46230252696668, Threshold: 0.03425741908792869  
Episode: 168, Score: -133.46685019536596, Threshold: 0.03357227070617012  
Episode: 169, Score: -10.405873155575078, Threshold: 0.032900825292046715  
Episode: 170, Score: -130.73948792134118, Threshold: 0.03224280878620578  
Episode: 171, Score: -133.1260746189509, Threshold: 0.031597952610481664  
Episode: 172, Score: -137.5130268195788, Threshold: 0.03096599355827203  
Episode: 173, Score: -389.9692146607478, Threshold: 0.030346673687106588  
Episode: 174, Score: -133.14319728026624, Threshold: 0.029739740213364455  
Episode: 175, Score: -130.9336013733314, Threshold: 0.029144945409097164  
Episode: 176, Score: -136.61884998434988, Threshold: 0.02856204650091522  
Episode: 177, Score: -239.80767563865206, Threshold: 0.027990805570896914  
Episode: 178, Score: -132.40751216976597, Threshold: 0.027430989459478975  
Episode: 179, Score: -374.5116754904423, Threshold: 0.026882369670289397  
Episode: 180, Score: -130.87214852159767, Threshold: 0.026344722276883607  
Episode: 181, Score: -131.8592619272187, Threshold: 0.025817827831345935  
Episode: 182, Score: -132.3795060384667, Threshold: 0.025301471274719014  
Episode: 183, Score: -133.99625561632956, Threshold: 0.024795441849224633  
Episode: 184, Score: -132.662574094201, Threshold: 0.02429953301224014  
Episode: 185, Score: -233.54479677227383, Threshold: 0.02381354235199534  
Episode: 186, Score: -136.9719440252845, Threshold: 0.02333727150495543  
Episode: 187, Score: -240.80452032547657, Threshold: 0.02287052607485632  
Episode: 188, Score: -322.82812901865987, Threshold: 0.022413115553359196  
Episode: 189, Score: -129.3671704694439, Threshold: 0.021964853242292012  
Episode: 190, Score: -10.074949935127844, Threshold: 0.02152555617744617  
Episode: 191, Score: -130.13594057806893, Threshold: 0.021095045053897248  
Episode: 192, Score: -6.89335763109835, Threshold: 0.020673144152819303  
Episode: 193, Score: -121.4961161453742, Threshold: 0.02025968126976292  
Episode: 194, Score: -378.61563388504476, Threshold: 0.01985448764436766  
Episode: 195, Score: -131.44959220452208, Threshold: 0.019457397891480308  
Episode: 196, Score: -121.31639563492867, Threshold: 0.0190682499336507  
Episode: 197, Score: -9.308486963051381, Threshold: 0.018686884934977686

Episode: 198, Score: -136.89162501149914, Threshold: 0.01831314723627813  
Episode: 199, Score: -271.59727745145705, Threshold: 0.017946884291552567  
Episode: 200, Score: -382.8961399102453, Threshold: 0.017587946605721515  
Episode: 201, Score: -126.95783951374226, Threshold: 0.017236187673607083  
Episode: 202, Score: -249.25494044539928, Threshold: 0.016891463920134943  
Episode: 203, Score: -124.90652249065955, Threshold: 0.016553634641732245  
Episode: 204, Score: -321.86474438437443, Threshold: 0.0162225619488976  
Episode: 205, Score: -129.6028208997155, Threshold: 0.015898110709919647  
Episode: 206, Score: -131.4730905536148, Threshold: 0.015580148495721254  
Episode: 207, Score: -413.7373619199539, Threshold: 0.015268545525806828  
Episode: 208, Score: -359.57779265401575, Threshold: 0.014963174615290691  
Episode: 209, Score: -246.1595087403537, Threshold: 0.014663911122984877  
Episode: 210, Score: -133.3730307903668, Threshold: 0.01437063290052518  
Episode: 211, Score: -240.7316228322308, Threshold: 0.014083220242514675  
Episode: 212, Score: -238.40380257710515, Threshold: 0.013801555837664382  
Episode: 213, Score: -129.37616726440726, Threshold: 0.013525524720911095  
Episode: 214, Score: -131.79996524559016, Threshold: 0.013255014226492874  
Episode: 215, Score: -228.3251346811995, Threshold: 0.012989913941963017  
Episode: 216, Score: -274.51197850491275, Threshold: 0.012730115663123756  
Episode: 217, Score: -124.01579599639277, Threshold: 0.012475513349861281  
Episode: 218, Score: -359.1739137914141, Threshold: 0.012226003082864055  
Episode: 219, Score: -135.70417217147897, Threshold: 0.011981483021206774  
Episode: 220, Score: -128.3492185086045, Threshold: 0.011741853360782638  
Episode: 221, Score: -253.41051684480567, Threshold: 0.011507016293566985  
Episode: 222, Score: -122.3550598639836, Threshold: 0.011276875967695645  
Episode: 223, Score: -246.6476068604252, Threshold: 0.011051338448341731  
Episode: 224, Score: -240.43939347808177, Threshold: 0.010830311679374895  
Episode: 225, Score: -243.01911429373737, Threshold: 0.010613705445787398  
Episode: 226, Score: -237.41150797699788, Threshold: 0.01040143133687165  
Episode: 227, Score: -10.407352802703272, Threshold: 0.010193402710134216  
Episode: 228, Score: -131.63005324806292, Threshold: 0.009989534655931531  
Episode: 229, Score: -133.2340166725071, Threshold: 0.0097897439628129  
Episode: 230, Score: -251.73496696910053, Threshold: 0.009593949083556642  
Episode: 231, Score: -257.87917888718687, Threshold: 0.009402070101885509  
Episode: 232, Score: -11.134336977969078, Threshold: 0.009214028699847799  
Episode: 233, Score: -130.193288422501, Threshold: 0.009029748125850843  
Episode: 234, Score: -359.40065453762344, Threshold: 0.008849153163333826  
Episode: 235, Score: -250.94401601954038, Threshold: 0.008672170100067149  
Episode: 236, Score: -131.21162799764232, Threshold: 0.008498726698065806  
Episode: 237, Score: -133.5837935278028, Threshold: 0.00832875216410449  
Episode: 238, Score: -243.36773895863556, Threshold: 0.0081621771208224  
Episode: 239, Score: -137.33209449979475, Threshold: 0.007998933578405951  
Episode: 240, Score: -134.40804327930672, Threshold: 0.007838954906837833  
Episode: 241, Score: -133.71621744050023, Threshold: 0.007682175808701076  
Episode: 242, Score: -128.99396278998952, Threshold: 0.007528532292527055  
Episode: 243, Score: -230.67227041611181, Threshold: 0.007377961646676513  
Episode: 244, Score: -134.99249279570472, Threshold: 0.007230402413742983  
Episode: 245, Score: -128.84654726250446, Threshold: 0.007085794365468123  
Episode: 246, Score: -249.55996242383435, Threshold: 0.00694407847815876  
Episode: 247, Score: -133.5228650960479, Threshold: 0.0068051969085955845  
Episode: 248, Score: -132.75860595937152, Threshold: 0.006669092970423673  
Episode: 249, Score: -126.33744894363127, Threshold: 0.00653571111015199  
Episode: 250, Score: -242.7858681589187, Threshold: 0.006404996888794895  
Episode: 251, Score: -129.9186917998027, Threshold: 0.006276896951018997  
Episode: 252, Score: -302.6567751696872, Threshold: 0.006151359011998617  
Episode: 253, Score: -124.1504631081348, Threshold: 0.006028331831758644  
Episode: 254, Score: -131.2500622346903, Threshold: 0.005907765195123471  
Episode: 255, Score: -129.06232614905807, Threshold: 0.005789609891221002  
Episode: 256, Score: -131.755136740485, Threshold: 0.005673817693396582  
Episode: 257, Score: -131.4197654459087, Threshold: 0.00556034133952865  
Episode: 258, Score: -244.86027920469235, Threshold: 0.005449134512738076  
Episode: 259, Score: -130.0113753763234, Threshold: 0.0053401518224833145  
Episode: 260, Score: -261.46399235395273, Threshold: 0.005233348786033648  
Episode: 261, Score: -233.25613812727602, Threshold: 0.005128681810312975  
Episode: 262, Score: -131.14757475933672, Threshold: 0.005026108174106716  
Episode: 263, Score: -128.42713420429595, Threshold: 0.004925586010624581

Episode: 264, Score: -126.79661729345275, Threshold: 0.004827074290412089  
Episode: 265, Score: -233.19457659995777, Threshold: 0.004730532804603847  
Episode: 266, Score: -125.17388395041446, Threshold: 0.00463592214851177  
Episode: 267, Score: -8.020088337079741, Threshold: 0.004543203705541535  
Episode: 268, Score: -133.25717309374338, Threshold: 0.004452339631430704  
Episode: 269, Score: -245.4619240356029, Threshold: 0.00436329283880209  
Episode: 270, Score: -130.13439099346627, Threshold: 0.004276026982026048  
Episode: 271, Score: -128.55438802711208, Threshold: 0.004190506442385527  
Episode: 272, Score: -253.19468116725847, Threshold: 0.004106696313537816  
Episode: 273, Score: -129.82402469521475, Threshold: 0.00402456238726706  
Episode: 274, Score: -119.33893792328965, Threshold: 0.003944071139521719  
Episode: 275, Score: -6.264191616268583, Threshold: 0.0038651897167312848  
Episode: 276, Score: -128.66211209530712, Threshold: 0.003787885922396659  
Episode: 277, Score: -130.40167413160313, Threshold: 0.003712128203948726  
Episode: 278, Score: -130.77631018467196, Threshold: 0.0036378856398697514  
Episode: 279, Score: -246.2584969629975, Threshold: 0.0035651279270723563  
Episode: 280, Score: -123.64212200754925, Threshold: 0.0034938253685309092  
Episode: 281, Score: -130.83597168283978, Threshold: 0.003423948861160291  
Episode: 282, Score: -7.250133434949519, Threshold: 0.0033554698839370852  
Episode: 283, Score: -8.817997354558482, Threshold: 0.0032883604862583436  
Episode: 284, Score: -6.632875769181768, Threshold: 0.0032225932765331767  
Episode: 285, Score: -409.68785767409867, Threshold: 0.003158141411002513  
Episode: 286, Score: -6.677609916190963, Threshold: 0.003094978582782463  
Episode: 287, Score: -129.43425538056582, Threshold: 0.0030330790111268136  
Episode: 288, Score: -393.49557632580445, Threshold: 0.0029724174309042773  
Episode: 289, Score: -129.7123348891668, Threshold: 0.002912969082286192  
Episode: 290, Score: -8.27240859711189, Threshold: 0.002854709700640468  
Episode: 291, Score: -129.65113839276734, Threshold: 0.0027976155066276585  
Episode: 292, Score: -124.45517231619542, Threshold: 0.0027416631964951054  
Episode: 293, Score: -243.70429649057212, Threshold: 0.0026868299325652034  
Episode: 294, Score: -248.01233794097854, Threshold: 0.002633093333913899  
Episode: 295, Score: -130.82128490122454, Threshold: 0.002580431467235621  
Episode: 296, Score: -6.611867468555783, Threshold: 0.0025288228378909086  
Episode: 297, Score: -122.2240873267804, Threshold: 0.0024782463811330903  
Episode: 298, Score: -131.0894097374712, Threshold: 0.0024286814535104282  
Episode: 299, Score: -130.62957733405153, Threshold: 0.0023801078244402194



Comparison

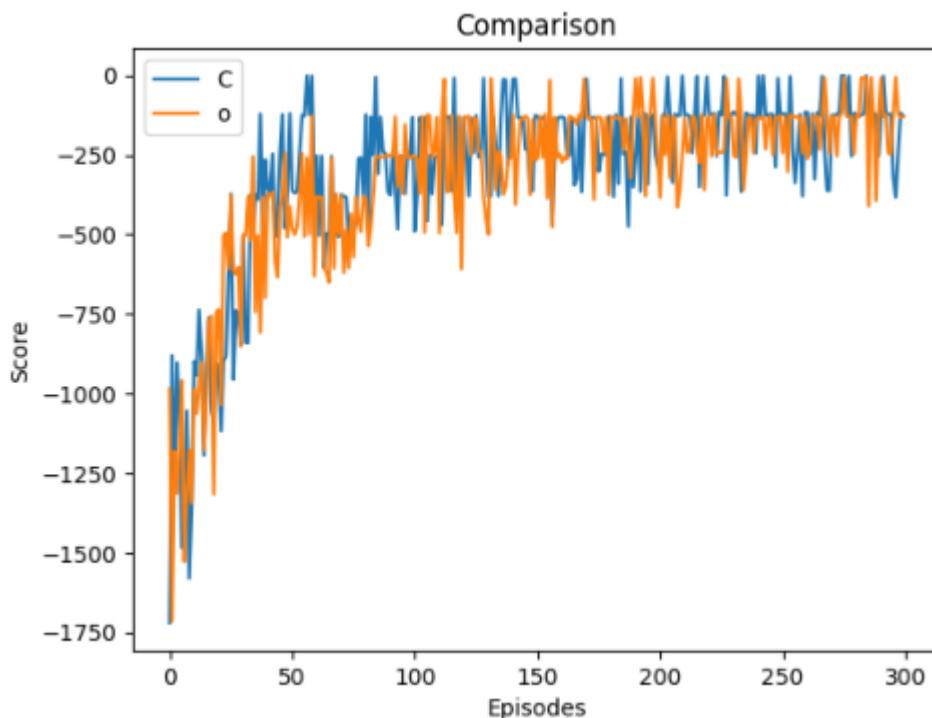


In [123...]: `evaluate_model(noisyAgent, 100)`

Evaluating Episode: 1, Score: -242.36706002028535  
Evaluating Episode: 2, Score: -121.79841189782114  
Evaluating Episode: 3, Score: -240.41279461565233  
Evaluating Episode: 4, Score: -127.72978182811494  
Evaluating Episode: 5, Score: -119.73980952027595  
Evaluating Episode: 6, Score: -242.78489675971537  
Evaluating Episode: 7, Score: -130.37836736195524  
Evaluating Episode: 8, Score: -122.83704554803361  
Evaluating Episode: 9, Score: -128.80260110947822  
Evaluating Episode: 10, Score: -128.36572473552474  
Evaluating Episode: 11, Score: -124.46357355136362  
Evaluating Episode: 12, Score: -128.7123786678771  
Evaluating Episode: 13, Score: -119.15434516381136  
Evaluating Episode: 14, Score: -118.20710338527971  
Evaluating Episode: 15, Score: -122.9499804636446  
Evaluating Episode: 16, Score: -132.32282399257068  
Evaluating Episode: 17, Score: -252.3042969684047  
Evaluating Episode: 18, Score: -242.06014793845304  
Evaluating Episode: 19, Score: -129.29443947002994  
Evaluating Episode: 20, Score: -126.67383153181504  
Evaluating Episode: 21, Score: -124.05232938443659  
Evaluating Episode: 22, Score: -240.48163891609556  
Evaluating Episode: 23, Score: -127.22346126487619  
Evaluating Episode: 24, Score: -123.41607989130762  
Evaluating Episode: 25, Score: -241.04745354948452  
Evaluating Episode: 26, Score: -235.33530177332167  
Evaluating Episode: 27, Score: -130.77644910114887  
Evaluating Episode: 28, Score: -117.91884420334218  
Evaluating Episode: 29, Score: -261.7378941436827  
Evaluating Episode: 30, Score: -10.65900671704475  
Evaluating Episode: 31, Score: -4.7690105203104745  
Evaluating Episode: 32, Score: -123.5923315608618  
Evaluating Episode: 33, Score: -128.79522410512146  
Evaluating Episode: 34, Score: -317.12173366472246  
Evaluating Episode: 35, Score: -122.53587416817477  
Evaluating Episode: 36, Score: -120.87126719761699  
Evaluating Episode: 37, Score: -123.50109332008114  
Evaluating Episode: 38, Score: -317.97586109478016  
Evaluating Episode: 39, Score: -129.95508612756615  
Evaluating Episode: 40, Score: -129.01812424541527  
Evaluating Episode: 41, Score: -4.999771415369745  
Evaluating Episode: 42, Score: -228.0391850188355  
Evaluating Episode: 43, Score: -247.04043867179928  
Evaluating Episode: 44, Score: -248.50175087987748  
Evaluating Episode: 45, Score: -126.06631938216445  
Evaluating Episode: 46, Score: -5.9490036478690635  
Evaluating Episode: 47, Score: -5.669579816471034  
Evaluating Episode: 48, Score: -127.14265740383054  
Evaluating Episode: 49, Score: -132.56695300866585  
Evaluating Episode: 50, Score: -228.82436658638403  
Evaluating Episode: 51, Score: -119.68595001060767  
Evaluating Episode: 52, Score: -124.90569421542075  
Evaluating Episode: 53, Score: -258.48083398729784  
Evaluating Episode: 54, Score: -340.95764186969194  
Evaluating Episode: 55, Score: -336.4574309867535  
Evaluating Episode: 56, Score: -4.7859963987506  
Evaluating Episode: 57, Score: -128.19325304296058  
Evaluating Episode: 58, Score: -4.662065502118951  
Evaluating Episode: 59, Score: -234.461348914293  
Evaluating Episode: 60, Score: -358.8138774331142  
Evaluating Episode: 61, Score: -128.70892855758814  
Evaluating Episode: 62, Score: -130.75532860578133  
Evaluating Episode: 63, Score: -121.00103329285785  
Evaluating Episode: 64, Score: -237.45303315501874  
Evaluating Episode: 65, Score: -376.45652234630296  
Evaluating Episode: 66, Score: -131.8435229159565

```
Evaluating Episode: 67, Score: -237.02629204065428
Evaluating Episode: 68, Score: -126.6774410594257
Evaluating Episode: 69, Score: -7.901125515898129
Evaluating Episode: 70, Score: -125.99050466155161
Evaluating Episode: 71, Score: -124.46279259255944
Evaluating Episode: 72, Score: -124.58370066246376
Evaluating Episode: 73, Score: -121.39449576586303
Evaluating Episode: 74, Score: -126.15911102288794
Evaluating Episode: 75, Score: -128.8333961592793
Evaluating Episode: 76, Score: -119.3833568998472
Evaluating Episode: 77, Score: -131.21257807976718
Evaluating Episode: 78, Score: -119.85755869736644
Evaluating Episode: 79, Score: -129.3097857009046
Evaluating Episode: 80, Score: -126.99079015085742
Evaluating Episode: 81, Score: -248.80541453029477
Evaluating Episode: 82, Score: -10.733780510478052
Evaluating Episode: 83, Score: -123.92133237665743
Evaluating Episode: 84, Score: -259.9016258284987
Evaluating Episode: 85, Score: -236.10531739797116
Evaluating Episode: 86, Score: -5.808687947021717
Evaluating Episode: 87, Score: -5.645634904386682
Evaluating Episode: 88, Score: -127.33127872752965
Evaluating Episode: 89, Score: -251.48364187553528
Evaluating Episode: 90, Score: -127.16749733798642
Evaluating Episode: 91, Score: -124.28845399144106
Evaluating Episode: 92, Score: -117.79334576643163
Evaluating Episode: 93, Score: -126.62748419583745
Evaluating Episode: 94, Score: -119.08132905322512
Evaluating Episode: 95, Score: -229.93901043990434
Evaluating Episode: 96, Score: -247.07964971131943
Evaluating Episode: 97, Score: -125.34302478988977
Evaluating Episode: 98, Score: -5.384091622532529
Evaluating Episode: 99, Score: -122.1778794820597
Evaluating Episode: 100, Score: -4.730923944923084
Final Score: -149.9369829998853
```

```
In [125...]: torch.save(noisyAgent.policy_net.state_dict(), "./rl_models/noisy_dqn.pt")
showImage("./Img_RL/noisy_ddqn.png")
```



Observations:

## Performance

- Noisy Netowrk performs slightly better than DDQN with a score of -149
- Training stability is around the same

### Training Time

- Surprisingly Noisy DQN didn't take that long to converge.

Overall, I think that adding the Noisy layer was beneficial as it allowed the model to explore the action space more and give better results.

## Evaluation of Final Model

```
In [126]: finalAgent, final_score = train_model_eps(NoisyDoubleDQNAgent, 500)
```

Episode: 0, Score: -1693.35497018581, Threshold: 1.0  
Episode: 1, Score: -1044.7013616337586, Threshold: 0.98  
Episode: 2, Score: -1187.431187324627, Threshold: 0.9603999999999999  
Episode: 3, Score: -1166.7596586114566, Threshold: 0.9411919999999999  
Episode: 4, Score: -1663.1783036962433, Threshold: 0.9223681599999999  
Episode: 5, Score: -950.4959029965835, Threshold: 0.9039207967999998  
Episode: 6, Score: -1680.7975025479834, Threshold: 0.8858423808639998  
Episode: 7, Score: -1100.980675995632, Threshold: 0.8681255332467198  
Episode: 8, Score: -1671.1550797556934, Threshold: 0.8507630225817854  
Episode: 9, Score: -974.6130274564158, Threshold: 0.8337477621301497  
Episode: 10, Score: -1305.3232484995503, Threshold: 0.8170728068875467  
Episode: 11, Score: -1072.0908032056427, Threshold: 0.8007313507497957  
Episode: 12, Score: -1361.8069988969464, Threshold: 0.7847167237347998  
Episode: 13, Score: -851.9698150368514, Threshold: 0.7690223892601038  
Episode: 14, Score: -1054.3927685195354, Threshold: 0.7536419414749017  
Episode: 15, Score: -1137.1286073679205, Threshold: 0.7385691026454037  
Episode: 16, Score: -1073.3082234401422, Threshold: 0.7237977205924956  
Episode: 17, Score: -985.8390572699592, Threshold: 0.7093217661806457  
Episode: 18, Score: -1021.3770201572906, Threshold: 0.6951353308570327  
Episode: 19, Score: -1088.295966057418, Threshold: 0.6812326242398921  
Episode: 20, Score: -756.2332162718371, Threshold: 0.6676079717550942  
Episode: 21, Score: -1017.8896736621604, Threshold: 0.6542558123199923  
Episode: 22, Score: -622.0508260226603, Threshold: 0.6411706960735924  
Episode: 23, Score: -839.5403833586398, Threshold: 0.6283472821521205  
Episode: 24, Score: -492.9209812680174, Threshold: 0.6157803365090782  
Episode: 25, Score: -994.757082294239, Threshold: 0.6034647297788965  
Episode: 26, Score: -254.9218431438629, Threshold: 0.5913954351833186  
Episode: 27, Score: -1031.5130283281892, Threshold: 0.5795675264796523  
Episode: 28, Score: -251.4691637480608, Threshold: 0.5679761759500592  
Episode: 29, Score: -622.0983251268722, Threshold: 0.5566166524310581  
Episode: 30, Score: -751.2666401102085, Threshold: 0.5454843193824369  
Episode: 31, Score: -250.4063085808503, Threshold: 0.5345746329947881  
Episode: 32, Score: -612.0949838580285, Threshold: 0.5238831403348924  
Episode: 33, Score: -379.42926431560574, Threshold: 0.5134054775281945  
Episode: 34, Score: -632.2597086739788, Threshold: 0.5031373679776306  
Episode: 35, Score: -318.95241355105117, Threshold: 0.493074620618078  
Episode: 36, Score: -253.9471485833062, Threshold: 0.48321312820571644  
Episode: 37, Score: -252.91026011903278, Threshold: 0.4735488656416021  
Episode: 38, Score: -495.81405681147305, Threshold: 0.46407788832877006  
Episode: 39, Score: -253.7527949385262, Threshold: 0.45479633056219465  
Episode: 40, Score: -253.11564808291172, Threshold: 0.44570040395095073  
Episode: 41, Score: -391.65974887530126, Threshold: 0.4367863958719317  
Episode: 42, Score: -370.08841459046926, Threshold: 0.42805066795449304  
Episode: 43, Score: -358.8497809902697, Threshold: 0.41948965459540316  
Episode: 44, Score: -378.26096487311383, Threshold: 0.4110998615034951  
Episode: 45, Score: -351.80838199120893, Threshold: 0.4028778642734252  
Episode: 46, Score: -242.4535547843022, Threshold: 0.39482030698795667  
Episode: 47, Score: -253.96015219202513, Threshold: 0.38692390084819756  
Episode: 48, Score: -493.1794910968926, Threshold: 0.3791854228312336  
Episode: 49, Score: -466.4630356083795, Threshold: 0.37160171437460887  
Episode: 50, Score: -375.9383282049181, Threshold: 0.3641696800871167  
Episode: 51, Score: -362.8171885889534, Threshold: 0.35688628648537435  
Episode: 52, Score: -277.1010251263711, Threshold: 0.34974856075566685  
Episode: 53, Score: -605.5754249593062, Threshold: 0.3427535895405535  
Episode: 54, Score: -374.72519415806335, Threshold: 0.33589851774974244  
Episode: 55, Score: -742.5115430635116, Threshold: 0.3291805473947476  
Episode: 56, Score: -375.7002257842152, Threshold: 0.32259693644685267  
Episode: 57, Score: -454.4327649451719, Threshold: 0.3161449977179156  
Episode: 58, Score: -610.4521836742341, Threshold: 0.3098220977635573  
Episode: 59, Score: -483.2991824094524, Threshold: 0.30362565580828615  
Episode: 60, Score: -469.24277076888575, Threshold: 0.2975531426921204  
Episode: 61, Score: -492.4389007279787, Threshold: 0.291602079838278  
Episode: 62, Score: -380.93332450851733, Threshold: 0.2857700382415124  
Episode: 63, Score: -615.4689209668416, Threshold: 0.2800546374766822  
Episode: 64, Score: -603.0458272575547, Threshold: 0.27445354472714856  
Episode: 65, Score: -380.86155325993514, Threshold: 0.2689644738326056

Episode: 66, Score: -506.6484861702787, Threshold: 0.26358518435595346  
Episode: 67, Score: -384.2418610822215, Threshold: 0.25831348066883436  
Episode: 68, Score: -478.29187071258076, Threshold: 0.2531472110554577  
Episode: 69, Score: -500.6481402834584, Threshold: 0.24808426683434853  
Episode: 70, Score: -374.78926237330217, Threshold: 0.24312258149766156  
Episode: 71, Score: -404.55769731197637, Threshold: 0.2382601298677083  
Episode: 72, Score: -377.49247173349823, Threshold: 0.23349492727035415  
Episode: 73, Score: -615.7763893892351, Threshold: 0.22882502872494706  
Episode: 74, Score: -498.84418647116075, Threshold: 0.22424852815044813  
Episode: 75, Score: -607.5121142464704, Threshold: 0.21976355758743915  
Episode: 76, Score: -379.5567381865368, Threshold: 0.21536828643569036  
Episode: 77, Score: -373.5880533728348, Threshold: 0.21106092070697655  
Episode: 78, Score: -129.1407829788933, Threshold: 0.20683970229283702  
Episode: 79, Score: -252.24653394837878, Threshold: 0.20270290824698028  
Episode: 80, Score: -229.57335259582018, Threshold: 0.19864885008204067  
Episode: 81, Score: -339.3916578821987, Threshold: 0.19467587308039985  
Episode: 82, Score: -121.3080997262096, Threshold: 0.19078235561879187  
Episode: 83, Score: -385.2261934408793, Threshold: 0.18696670850641603  
Episode: 84, Score: -374.529928514752, Threshold: 0.18322737433628772  
Episode: 85, Score: -601.8918953910307, Threshold: 0.17956282684956196  
Episode: 86, Score: -481.5046133265137, Threshold: 0.1759715703125707  
Episode: 87, Score: -261.68214968828914, Threshold: 0.1724521389063193  
Episode: 88, Score: -257.9762026196427, Threshold: 0.16900309612819292  
Episode: 89, Score: -371.2819231546712, Threshold: 0.16562303420562907  
Episode: 90, Score: -255.874249248905, Threshold: 0.16231057352151648  
Episode: 91, Score: -368.22323621699, Threshold: 0.15906436205108615  
Episode: 92, Score: -251.47573208103788, Threshold: 0.15588307481006441  
Episode: 93, Score: -482.3163371595018, Threshold: 0.15276541331386312  
Episode: 94, Score: -255.3629195821601, Threshold: 0.14971010504758586  
Episode: 95, Score: -378.23023184379286, Threshold: 0.14671590294663414  
Episode: 96, Score: -129.50984822344526, Threshold: 0.14378158488770146  
Episode: 97, Score: -121.54664113293168, Threshold: 0.14090595318994742  
Episode: 98, Score: -130.36533102164572, Threshold: 0.13808783412614847  
Episode: 99, Score: -243.93517006870584, Threshold: 0.13532607744362551  
Episode: 100, Score: -239.316807294195, Threshold: 0.132619555894753  
Episode: 101, Score: -254.3327610734647, Threshold: 0.12996716477685794  
Episode: 102, Score: -4.44548908733229, Threshold: 0.12736782148132078  
Episode: 103, Score: -124.51991017264791, Threshold: 0.12482046505169436  
Episode: 104, Score: -120.92523981961476, Threshold: 0.12232405575066048  
Episode: 105, Score: -130.00641157461587, Threshold: 0.11987757463564727  
Episode: 106, Score: -5.012421201326622, Threshold: 0.11748002314293432  
Episode: 107, Score: -16.24120621300747, Threshold: 0.11513042268007563  
Episode: 108, Score: -127.78474629651828, Threshold: 0.11282781422647412  
Episode: 109, Score: -505.60036735218455, Threshold: 0.11057125794194463  
Episode: 110, Score: -129.08812718200107, Threshold: 0.10835983278310574  
Episode: 111, Score: -123.94125539343224, Threshold: 0.10619263612744362  
Episode: 112, Score: -238.02994577922877, Threshold: 0.10406878340489474  
Episode: 113, Score: -249.3031409348078, Threshold: 0.10198740773679685  
Episode: 114, Score: -121.71533080423409, Threshold: 0.09994765958206091  
Episode: 115, Score: -128.8016862549981, Threshold: 0.0979487063904197  
Episode: 116, Score: -133.8312657310211, Threshold: 0.0959897322626113  
Episode: 117, Score: -504.52764852931705, Threshold: 0.09406993761735907  
Episode: 118, Score: -127.65127683968556, Threshold: 0.0921885388650119  
Episode: 119, Score: -378.189710630233, Threshold: 0.09034476808771166  
Episode: 120, Score: -134.17872592156985, Threshold: 0.08853787272595742  
Episode: 121, Score: -381.24830775564396, Threshold: 0.08676711527143827  
Episode: 122, Score: -258.96732473181856, Threshold: 0.0850317729660095  
Episode: 123, Score: -10.047065846190371, Threshold: 0.08333113750668932  
Episode: 124, Score: -252.56865548710462, Threshold: 0.08166451475655553  
Episode: 125, Score: -254.38025664675897, Threshold: 0.08003122446142442  
Episode: 126, Score: -11.770744878202288, Threshold: 0.07843059997219594  
Episode: 127, Score: -253.63673292857155, Threshold: 0.07686198797275202  
Episode: 128, Score: -252.49789198063243, Threshold: 0.07532474821329699  
Episode: 129, Score: -249.0191269578593, Threshold: 0.07381825324903105  
Episode: 130, Score: -374.4862939688033, Threshold: 0.07234188818405042  
Episode: 131, Score: -133.1430749091091, Threshold: 0.07089505042036941

Episode: 132, Score: -255.82832331525086, Threshold: 0.06947714941196202  
Episode: 133, Score: -243.4262652041952, Threshold: 0.06808760642372277  
Episode: 134, Score: -9.577236691760023, Threshold: 0.06672585429524831  
Episode: 135, Score: -380.57773298061187, Threshold: 0.06539133720934334  
Episode: 136, Score: -381.8370188064712, Threshold: 0.06408351046515648  
Episode: 137, Score: -484.6195353618147, Threshold: 0.06280184025585335  
Episode: 138, Score: -343.03276146633715, Threshold: 0.061545803450736285  
Episode: 139, Score: -243.00896735469937, Threshold: 0.060314887381721555  
Episode: 140, Score: -8.231707534335094, Threshold: 0.05910858963408712  
Episode: 141, Score: -236.3455442214715, Threshold: 0.05792641784140538  
Episode: 142, Score: -121.43567029841816, Threshold: 0.05676788948457727  
Episode: 143, Score: -355.55339096771155, Threshold: 0.055632531694885724  
Episode: 144, Score: -249.52229921070068, Threshold: 0.054519881060988006  
Episode: 145, Score: -267.3445258041269, Threshold: 0.05342948343976824  
Episode: 146, Score: -9.060281514806315, Threshold: 0.052360893770972874  
Episode: 147, Score: -349.2783001547102, Threshold: 0.05131367589555342  
Episode: 148, Score: -238.54905565042696, Threshold: 0.05028740237764235  
Episode: 149, Score: -249.64178832142014, Threshold: 0.0492816543300895  
Episode: 150, Score: -230.0731296115394, Threshold: 0.04829602124348771  
Episode: 151, Score: -6.487825649092408, Threshold: 0.04733010081861796  
Episode: 152, Score: -127.67189082427957, Threshold: 0.0463834988022456  
Episode: 153, Score: -133.3965423906245, Threshold: 0.04545582882620069  
Episode: 154, Score: -124.54475700965438, Threshold: 0.04454671224967667  
Episode: 155, Score: -234.4005682860191, Threshold: 0.043655778004683135  
Episode: 156, Score: -124.303786903477, Threshold: 0.04278266244458947  
Episode: 157, Score: -125.81087278490274, Threshold: 0.04192700919569768  
Episode: 158, Score: -435.8590540841627, Threshold: 0.04108846901178373  
Episode: 159, Score: -355.1111101766543, Threshold: 0.04026669963154805  
Episode: 160, Score: -134.03783676066632, Threshold: 0.03946136563891709  
Episode: 161, Score: -128.26963699839985, Threshold: 0.03867213832613875  
Episode: 162, Score: -511.49887940154525, Threshold: 0.037898695559615975  
Episode: 163, Score: -133.59076647374195, Threshold: 0.03714072164842366  
Episode: 164, Score: -7.065659705204014, Threshold: 0.03639790721545518  
Episode: 165, Score: -5.8125016472078315, Threshold: 0.035669949071146075  
Episode: 166, Score: -241.80882672020448, Threshold: 0.034956550089723155  
Episode: 167, Score: -6.130434694336028, Threshold: 0.03425741908792869  
Episode: 168, Score: -244.32143493338998, Threshold: 0.03357227070617012  
Episode: 169, Score: -249.80288217593304, Threshold: 0.032900825292046715  
Episode: 170, Score: -6.617801629744868, Threshold: 0.03224280878620578  
Episode: 171, Score: -127.51036412287941, Threshold: 0.031597952610481664  
Episode: 172, Score: -129.68300342228864, Threshold: 0.03096599355827203  
Episode: 173, Score: -129.87688466367314, Threshold: 0.030346673687106588  
Episode: 174, Score: -258.9697250933714, Threshold: 0.029739740213364455  
Episode: 175, Score: -247.11579757720327, Threshold: 0.029144945409097164  
Episode: 176, Score: -5.517044935908453, Threshold: 0.02856204650091522  
Episode: 177, Score: -7.520298120614858, Threshold: 0.027990805570896914  
Episode: 178, Score: -5.820944658031889, Threshold: 0.027430989459478975  
Episode: 179, Score: -373.5793770903151, Threshold: 0.026882369670289397  
Episode: 180, Score: -120.7584297324006, Threshold: 0.026344722276883607  
Episode: 181, Score: -284.44517009018944, Threshold: 0.025817827831345935  
Episode: 182, Score: -128.09973154202038, Threshold: 0.025301471274719014  
Episode: 183, Score: -131.67421167644684, Threshold: 0.024795441849224633  
Episode: 184, Score: -241.58582813737766, Threshold: 0.02429953301224014  
Episode: 185, Score: -119.32884126963016, Threshold: 0.02381354235199534  
Episode: 186, Score: -127.161903199196, Threshold: 0.02333727150495543  
Episode: 187, Score: -252.2461746745347, Threshold: 0.02287052607485632  
Episode: 188, Score: -257.78772692067963, Threshold: 0.022413115553359196  
Episode: 189, Score: -8.915555941266089, Threshold: 0.021964853242292012  
Episode: 190, Score: -132.05240835003193, Threshold: 0.02152555617744617  
Episode: 191, Score: -132.05328424461183, Threshold: 0.021095045053897248  
Episode: 192, Score: -125.82504634505429, Threshold: 0.020673144152819303  
Episode: 193, Score: -7.8742635317732175, Threshold: 0.02025968126976292  
Episode: 194, Score: -127.95912375818102, Threshold: 0.01985448764436766  
Episode: 195, Score: -8.561941464800697, Threshold: 0.019457397891480308  
Episode: 196, Score: -241.60900956102748, Threshold: 0.0190682499336507  
Episode: 197, Score: -255.56685009117152, Threshold: 0.018686884934977686

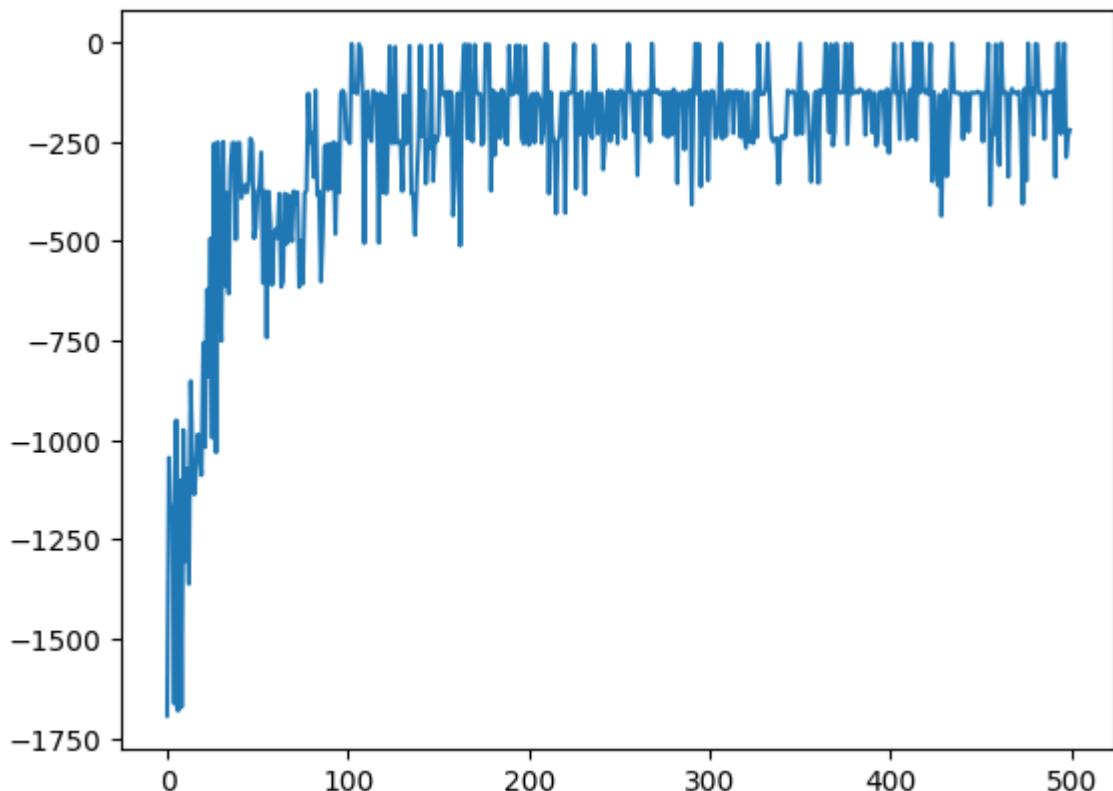
Episode: 198, Score: -10.421464345486479, Threshold: 0.01831314723627813  
Episode: 199, Score: -245.10133955997702, Threshold: 0.017946884291552567  
Episode: 200, Score: -259.0325616838892, Threshold: 0.017587946605721515  
Episode: 201, Score: -136.68042772508258, Threshold: 0.017236187673607083  
Episode: 202, Score: -129.2997416190854, Threshold: 0.016891463920134943  
Episode: 203, Score: -253.11517099629927, Threshold: 0.016553634641732245  
Episode: 204, Score: -128.15639698501215, Threshold: 0.0162225619488976  
Episode: 205, Score: -131.52386203350716, Threshold: 0.015898110709919647  
Episode: 206, Score: -143.2862949619459, Threshold: 0.015580148495721254  
Episode: 207, Score: -253.88748233570038, Threshold: 0.015268545525806828  
Episode: 208, Score: -227.85564618912457, Threshold: 0.014963174615290691  
Episode: 209, Score: -5.099623849586467, Threshold: 0.014663911122984877  
Episode: 210, Score: -6.760358334152624, Threshold: 0.01437063290052518  
Episode: 211, Score: -380.75519808228245, Threshold: 0.014083220242514675  
Episode: 212, Score: -125.55545198564596, Threshold: 0.013801555837664382  
Episode: 213, Score: -232.25940446150796, Threshold: 0.013525524720911095  
Episode: 214, Score: -254.0633553762833, Threshold: 0.013255014226492874  
Episode: 215, Score: -430.9086101261318, Threshold: 0.012989913941963017  
Episode: 216, Score: -251.8634621759229, Threshold: 0.012730115663123756  
Episode: 217, Score: -235.83149006202217, Threshold: 0.012475513349861281  
Episode: 218, Score: -127.57769059011346, Threshold: 0.012226003082864055  
Episode: 219, Score: -129.0751967459955, Threshold: 0.011981483021206774  
Episode: 220, Score: -429.4250039516096, Threshold: 0.011741853360782638  
Episode: 221, Score: -128.4915795820303, Threshold: 0.011507016293566985  
Episode: 222, Score: -126.68769216244127, Threshold: 0.011276875967695645  
Episode: 223, Score: -128.91205267610326, Threshold: 0.011051338448341731  
Episode: 224, Score: -121.7697507386954, Threshold: 0.010830311679374895  
Episode: 225, Score: -4.738792165354284, Threshold: 0.010613705445787398  
Episode: 226, Score: -367.7540867658613, Threshold: 0.01040143133687165  
Episode: 227, Score: -131.21090614871898, Threshold: 0.010193402710134216  
Episode: 228, Score: -232.1535044241282, Threshold: 0.009989534655931531  
Episode: 229, Score: -124.37911777637022, Threshold: 0.0097897439628129  
Episode: 230, Score: -229.2847436549556, Threshold: 0.009593949083556642  
Episode: 231, Score: -382.4583222448923, Threshold: 0.009402070101885509  
Episode: 232, Score: -129.54533516470303, Threshold: 0.009214028699847799  
Episode: 233, Score: -127.76035215329917, Threshold: 0.009029748125850843  
Episode: 234, Score: -230.63001303160055, Threshold: 0.008849153163333826  
Episode: 235, Score: -241.6319953194891, Threshold: 0.008672170100067149  
Episode: 236, Score: -7.178303875804585, Threshold: 0.008498726698065806  
Episode: 237, Score: -128.44356212806292, Threshold: 0.00832875216410449  
Episode: 238, Score: -230.6134239287255, Threshold: 0.0081621771208224  
Episode: 239, Score: -130.907802691544, Threshold: 0.007998933578405951  
Episode: 240, Score: -128.9227124486095, Threshold: 0.007838954906837833  
Episode: 241, Score: -320.1052166553501, Threshold: 0.007682175808701076  
Episode: 242, Score: -251.8656246541552, Threshold: 0.007528532292527055  
Episode: 243, Score: -136.70366932105526, Threshold: 0.007377961646676513  
Episode: 244, Score: -121.49853477953526, Threshold: 0.007230402413742983  
Episode: 245, Score: -249.0778656038555, Threshold: 0.007085794365468123  
Episode: 246, Score: -122.76526246718043, Threshold: 0.00694407847815876  
Episode: 247, Score: -226.891021764672, Threshold: 0.0068051969085955845  
Episode: 248, Score: -126.32573355721978, Threshold: 0.006669092970423673  
Episode: 249, Score: -254.75893271343247, Threshold: 0.00653571111015199  
Episode: 250, Score: -124.32399333522939, Threshold: 0.006404996888794895  
Episode: 251, Score: -121.90285524535966, Threshold: 0.006276896951018997  
Episode: 252, Score: -129.17848314594065, Threshold: 0.006151359011998617  
Episode: 253, Score: -242.08718764530263, Threshold: 0.006028331831758644  
Episode: 254, Score: -128.11920468076244, Threshold: 0.005907765195123471  
Episode: 255, Score: -3.7341909382684255, Threshold: 0.005789609891221002  
Episode: 256, Score: -124.81983661344096, Threshold: 0.005673817693396582  
Episode: 257, Score: -124.41053462355585, Threshold: 0.00556034133952865  
Episode: 258, Score: -224.25773347055215, Threshold: 0.005449134512738076  
Episode: 259, Score: -128.80998636126117, Threshold: 0.0053401518224833145  
Episode: 260, Score: -334.95717575842957, Threshold: 0.005233348786033648  
Episode: 261, Score: -123.28274837365777, Threshold: 0.005128681810312975  
Episode: 262, Score: -128.21657797500342, Threshold: 0.005026108174106716  
Episode: 263, Score: -130.35256349900143, Threshold: 0.004925586010624581

Episode: 264, Score: -126.58866544522606, Threshold: 0.004827074290412089  
Episode: 265, Score: -233.64126220600295, Threshold: 0.004730532804603847  
Episode: 266, Score: -122.98364741770666, Threshold: 0.00463592214851177  
Episode: 267, Score: -249.98149921465898, Threshold: 0.004543203705541535  
Episode: 268, Score: -3.6440197234843756, Threshold: 0.004452339631430704  
Episode: 269, Score: -123.25635208510164, Threshold: 0.00436329283880209  
Episode: 270, Score: -126.02182582026157, Threshold: 0.004276026982026048  
Episode: 271, Score: -119.81469070472379, Threshold: 0.004190506442385527  
Episode: 272, Score: -130.97173165189224, Threshold: 0.004106696313537816  
Episode: 273, Score: -129.0848503031083, Threshold: 0.00402456238726706  
Episode: 274, Score: -126.05341362014612, Threshold: 0.003944071139521719  
Episode: 275, Score: -234.19778139910858, Threshold: 0.0038651897167312848  
Episode: 276, Score: -126.05121383813388, Threshold: 0.003787885922396659  
Episode: 277, Score: -238.93658413344923, Threshold: 0.003712128203948726  
Episode: 278, Score: -125.65910379074644, Threshold: 0.0036378856398697514  
Episode: 279, Score: -232.3413953425864, Threshold: 0.0035651279270723563  
Episode: 280, Score: -118.61497640402466, Threshold: 0.0034938253685309092  
Episode: 281, Score: -127.1659228299435, Threshold: 0.003423948861160291  
Episode: 282, Score: -354.704053126571, Threshold: 0.0033554698839370852  
Episode: 283, Score: -129.13070721645354, Threshold: 0.0032883604862583436  
Episode: 284, Score: -133.46699188359702, Threshold: 0.0032225932765331767  
Episode: 285, Score: -124.9937056004827, Threshold: 0.003158141411002513  
Episode: 286, Score: -270.0368050476401, Threshold: 0.003094978582782463  
Episode: 287, Score: -127.6664639473969, Threshold: 0.0030330790111268136  
Episode: 288, Score: -125.53929757348443, Threshold: 0.0029724174309042773  
Episode: 289, Score: -120.67155122246301, Threshold: 0.002912969082286192  
Episode: 290, Score: -408.01478233067695, Threshold: 0.002854709700640468  
Episode: 291, Score: -129.62574175344656, Threshold: 0.0027976155066276585  
Episode: 292, Score: -4.122608961323103, Threshold: 0.0027416631964951054  
Episode: 293, Score: -129.40836586272155, Threshold: 0.0026868299325652034  
Episode: 294, Score: -4.424298777959366, Threshold: 0.002633093333913899  
Episode: 295, Score: -362.95427916019526, Threshold: 0.002580431467235621  
Episode: 296, Score: -128.18526172419342, Threshold: 0.0025288228378909086  
Episode: 297, Score: -231.85545082917122, Threshold: 0.0024782463811330903  
Episode: 298, Score: -124.5620320417389, Threshold: 0.0024286814535104282  
Episode: 299, Score: -347.47208339985383, Threshold: 0.0023801078244402194  
Episode: 300, Score: -128.63283277727956, Threshold: 0.002332505667951415  
Episode: 301, Score: -122.71576179188766, Threshold: 0.0022858555545923866  
Episode: 302, Score: -124.01195780281554, Threshold: 0.002240138443500539  
Episode: 303, Score: -240.5805509476954, Threshold: 0.002195335674630528  
Episode: 304, Score: -234.77249104659163, Threshold: 0.0021514289611379175  
Episode: 305, Score: -128.02035545077644, Threshold: 0.002108400381915159  
Episode: 306, Score: -3.0092350870321285, Threshold: 0.002066232374276856  
Episode: 307, Score: -242.97357455148583, Threshold: 0.002024907726791319  
Episode: 308, Score: -127.62582634355302, Threshold: 0.0019844095722554925  
Episode: 309, Score: -235.89630528137013, Threshold: 0.0019447213808103825  
Episode: 310, Score: -122.14682274145682, Threshold: 0.0019058269531941749  
Episode: 311, Score: -125.04450777891005, Threshold: 0.0018677104141302913  
Episode: 312, Score: -124.10374374835375, Threshold: 0.0018303562058476855  
Episode: 313, Score: -228.93406621663866, Threshold: 0.0017937490817307316  
Episode: 314, Score: -123.18911585879493, Threshold: 0.001757874100096117  
Episode: 315, Score: -122.99380491646208, Threshold: 0.0017227166180941945  
Episode: 316, Score: -125.51422718092286, Threshold: 0.0016882622857323107  
Episode: 317, Score: -238.74775658304713, Threshold: 0.0016544970400176645  
Episode: 318, Score: -123.62982422813482, Threshold: 0.0016214070992173113  
Episode: 319, Score: -236.31671880521338, Threshold: 0.001588978957232965  
Episode: 320, Score: -265.54874511244844, Threshold: 0.0015571993780883057  
Episode: 321, Score: -127.88407206343378, Threshold: 0.0015260553905265396  
Episode: 322, Score: -250.34540338770503, Threshold: 0.0014955342827160088  
Episode: 323, Score: -229.5105550075582, Threshold: 0.0014656235970616886  
Episode: 324, Score: -254.0963151769363, Threshold: 0.0014363111251204548  
Episode: 325, Score: -123.06673080067583, Threshold: 0.0014075849026180458  
Episode: 326, Score: -239.1616450603387, Threshold: 0.001379433204565685  
Episode: 327, Score: -5.539390927624812, Threshold: 0.0013518445404743713  
Episode: 328, Score: -131.51195868783958, Threshold: 0.0013248076496648838  
Episode: 329, Score: -130.78818987803297, Threshold: 0.0012983114966715861

Episode: 330, Score: -130.07549295703978, Threshold: 0.0012723452667381544  
Episode: 331, Score: -116.01240441763463, Threshold: 0.0012468983614033913  
Episode: 332, Score: -2.7950780509626196, Threshold: 0.0012219603941753235  
Episode: 333, Score: -126.92737040758983, Threshold: 0.001197521186291817  
Episode: 334, Score: -236.6550747672987, Threshold: 0.0011735707625659807  
Episode: 335, Score: -249.9124199514705, Threshold: 0.001150099347314661  
Episode: 336, Score: -243.68946771269202, Threshold: 0.0011270973603683678  
Episode: 337, Score: -241.42870004889025, Threshold: 0.0011045554131610005  
Episode: 338, Score: -355.61751528620164, Threshold: 0.0010824643048977804  
Episode: 339, Score: -244.14977438424123, Threshold: 0.0010608150187998249  
Episode: 340, Score: -235.16868065364258, Threshold: 0.0010395987184238283  
Episode: 341, Score: -242.77203047201434, Threshold: 0.0010188067440553518  
Episode: 342, Score: -236.83111835830064, Threshold: 0.0009984306091742447  
Episode: 343, Score: -121.22629243364277, Threshold: 0.0009984306091742447  
Episode: 344, Score: -128.3438994630502, Threshold: 0.0009984306091742447  
Episode: 345, Score: -126.63672223430719, Threshold: 0.0009984306091742447  
Episode: 346, Score: -122.76944138699035, Threshold: 0.0009984306091742447  
Episode: 347, Score: -132.1709950281189, Threshold: 0.0009984306091742447  
Episode: 348, Score: -233.11851337080424, Threshold: 0.0009984306091742447  
Episode: 349, Score: -119.54806616659842, Threshold: 0.0009984306091742447  
Episode: 350, Score: -3.2131211838498097, Threshold: 0.0009984306091742447  
Episode: 351, Score: -230.78214403927677, Threshold: 0.0009984306091742447  
Episode: 352, Score: -127.37968163497304, Threshold: 0.0009984306091742447  
Episode: 353, Score: -126.09150707463553, Threshold: 0.0009984306091742447  
Episode: 354, Score: -133.16541853379053, Threshold: 0.0009984306091742447  
Episode: 355, Score: -242.91891354085985, Threshold: 0.0009984306091742447  
Episode: 356, Score: -351.09190469824216, Threshold: 0.0009984306091742447  
Episode: 357, Score: -254.0441768930438, Threshold: 0.0009984306091742447  
Episode: 358, Score: -128.4664655502076, Threshold: 0.0009984306091742447  
Episode: 359, Score: -124.05173588454196, Threshold: 0.0009984306091742447  
Episode: 360, Score: -354.2172076745065, Threshold: 0.0009984306091742447  
Episode: 361, Score: -119.72042632111193, Threshold: 0.0009984306091742447  
Episode: 362, Score: -128.151275510666, Threshold: 0.0009984306091742447  
Episode: 363, Score: -126.67827115547477, Threshold: 0.0009984306091742447  
Episode: 364, Score: -3.727482172846356, Threshold: 0.0009984306091742447  
Episode: 365, Score: -124.39447982431257, Threshold: 0.0009984306091742447  
Episode: 366, Score: -227.10257056966316, Threshold: 0.0009984306091742447  
Episode: 367, Score: -5.533854069200426, Threshold: 0.0009984306091742447  
Episode: 368, Score: -260.1527833548711, Threshold: 0.0009984306091742447  
Episode: 369, Score: -226.39129131902268, Threshold: 0.0009984306091742447  
Episode: 370, Score: -3.596901555704534, Threshold: 0.0009984306091742447  
Episode: 371, Score: -119.40661748539421, Threshold: 0.0009984306091742447  
Episode: 372, Score: -127.77661998549604, Threshold: 0.0009984306091742447  
Episode: 373, Score: -126.39454920002204, Threshold: 0.0009984306091742447  
Episode: 374, Score: -125.27837550661364, Threshold: 0.0009984306091742447  
Episode: 375, Score: -4.393760164259168, Threshold: 0.0009984306091742447  
Episode: 376, Score: -256.38418762831793, Threshold: 0.0009984306091742447  
Episode: 377, Score: -128.16665675702376, Threshold: 0.0009984306091742447  
Episode: 378, Score: -3.5231902016119627, Threshold: 0.0009984306091742447  
Episode: 379, Score: -129.8297750902185, Threshold: 0.0009984306091742447  
Episode: 380, Score: -124.95710383276766, Threshold: 0.0009984306091742447  
Episode: 381, Score: -123.06979775168236, Threshold: 0.0009984306091742447  
Episode: 382, Score: -127.65535369710838, Threshold: 0.0009984306091742447  
Episode: 383, Score: -117.92100574456528, Threshold: 0.0009984306091742447  
Episode: 384, Score: -122.28062312851894, Threshold: 0.0009984306091742447  
Episode: 385, Score: -125.50196335944129, Threshold: 0.0009984306091742447  
Episode: 386, Score: -233.49670636967525, Threshold: 0.0009984306091742447  
Episode: 387, Score: -126.82029044542628, Threshold: 0.0009984306091742447  
Episode: 388, Score: -121.64547653236988, Threshold: 0.0009984306091742447  
Episode: 389, Score: -126.72086550627556, Threshold: 0.0009984306091742447  
Episode: 390, Score: -227.70654501798836, Threshold: 0.0009984306091742447  
Episode: 391, Score: -127.87259865686111, Threshold: 0.0009984306091742447  
Episode: 392, Score: -260.5958460037028, Threshold: 0.0009984306091742447  
Episode: 393, Score: -233.90448280979516, Threshold: 0.0009984306091742447  
Episode: 394, Score: -126.21346499670638, Threshold: 0.0009984306091742447  
Episode: 395, Score: -124.05199398536797, Threshold: 0.0009984306091742447

Episode: 396, Score: -129.19185380690698, Threshold: 0.0009984306091742447  
Episode: 397, Score: -256.81546965219115, Threshold: 0.0009984306091742447  
Episode: 398, Score: -117.65038020911719, Threshold: 0.0009984306091742447  
Episode: 399, Score: -278.800106152612, Threshold: 0.0009984306091742447  
Episode: 400, Score: -131.2587882477375, Threshold: 0.0009984306091742447  
Episode: 401, Score: -120.42011955843158, Threshold: 0.0009984306091742447  
Episode: 402, Score: -4.1981151357968916, Threshold: 0.0009984306091742447  
Episode: 403, Score: -129.70461730281886, Threshold: 0.0009984306091742447  
Episode: 404, Score: -126.09749655327369, Threshold: 0.0009984306091742447  
Episode: 405, Score: -244.18411478688168, Threshold: 0.0009984306091742447  
Episode: 406, Score: -4.417676048736228, Threshold: 0.0009984306091742447  
Episode: 407, Score: -122.08130442633092, Threshold: 0.0009984306091742447  
Episode: 408, Score: -128.07059068302078, Threshold: 0.0009984306091742447  
Episode: 409, Score: -244.6843858844591, Threshold: 0.0009984306091742447  
Episode: 410, Score: -240.66940637556365, Threshold: 0.0009984306091742447  
Episode: 411, Score: -118.92724335555774, Threshold: 0.0009984306091742447  
Episode: 412, Score: -237.21111849087055, Threshold: 0.0009984306091742447  
Episode: 413, Score: -2.3285616336016717, Threshold: 0.0009984306091742447  
Episode: 414, Score: -245.97715599545913, Threshold: 0.0009984306091742447  
Episode: 415, Score: -3.273770363777428, Threshold: 0.0009984306091742447  
Episode: 416, Score: -127.61101141962033, Threshold: 0.0009984306091742447  
Episode: 417, Score: -3.115853264605221, Threshold: 0.0009984306091742447  
Episode: 418, Score: -119.50502838742935, Threshold: 0.0009984306091742447  
Episode: 419, Score: -127.85024118545226, Threshold: 0.0009984306091742447  
Episode: 420, Score: -119.73921047412172, Threshold: 0.0009984306091742447  
Episode: 421, Score: -131.526445233116, Threshold: 0.0009984306091742447  
Episode: 422, Score: -4.808007013892123, Threshold: 0.0009984306091742447  
Episode: 423, Score: -349.8160119075333, Threshold: 0.0009984306091742447  
Episode: 424, Score: -125.13610360775013, Threshold: 0.0009984306091742447  
Episode: 425, Score: -249.84231344887186, Threshold: 0.0009984306091742447  
Episode: 426, Score: -360.14285616929965, Threshold: 0.0009984306091742447  
Episode: 427, Score: -132.2490389604574, Threshold: 0.0009984306091742447  
Episode: 428, Score: -436.9239865104846, Threshold: 0.0009984306091742447  
Episode: 429, Score: -234.50757265803804, Threshold: 0.0009984306091742447  
Episode: 430, Score: -120.57523309224044, Threshold: 0.0009984306091742447  
Episode: 431, Score: -337.7825686120756, Threshold: 0.0009984306091742447  
Episode: 432, Score: -226.21041062694007, Threshold: 0.0009984306091742447  
Episode: 433, Score: -128.87370907665493, Threshold: 0.0009984306091742447  
Episode: 434, Score: -3.32034988085632, Threshold: 0.0009984306091742447  
Episode: 435, Score: -124.86986158589363, Threshold: 0.0009984306091742447  
Episode: 436, Score: -127.73002189668455, Threshold: 0.0009984306091742447  
Episode: 437, Score: -128.292014493034, Threshold: 0.0009984306091742447  
Episode: 438, Score: -127.19008218987626, Threshold: 0.0009984306091742447  
Episode: 439, Score: -131.96147904970988, Threshold: 0.0009984306091742447  
Episode: 440, Score: -243.9487271516827, Threshold: 0.0009984306091742447  
Episode: 441, Score: -127.56030183946807, Threshold: 0.0009984306091742447  
Episode: 442, Score: -128.97980995327563, Threshold: 0.0009984306091742447  
Episode: 443, Score: -225.55036832429158, Threshold: 0.0009984306091742447  
Episode: 444, Score: -127.93102413696509, Threshold: 0.0009984306091742447  
Episode: 445, Score: -126.93918931108621, Threshold: 0.0009984306091742447  
Episode: 446, Score: -127.88220524288589, Threshold: 0.0009984306091742447  
Episode: 447, Score: -128.442360080392, Threshold: 0.0009984306091742447  
Episode: 448, Score: -123.98963135998424, Threshold: 0.0009984306091742447  
Episode: 449, Score: -127.20703613241737, Threshold: 0.0009984306091742447  
Episode: 450, Score: -127.30762548971347, Threshold: 0.0009984306091742447  
Episode: 451, Score: -249.6984466277028, Threshold: 0.0009984306091742447  
Episode: 452, Score: -128.50082943163082, Threshold: 0.0009984306091742447  
Episode: 453, Score: -123.9494411738365, Threshold: 0.0009984306091742447  
Episode: 454, Score: -3.661498113283917, Threshold: 0.0009984306091742447  
Episode: 455, Score: -409.5245428965399, Threshold: 0.0009984306091742447  
Episode: 456, Score: -222.64904553744745, Threshold: 0.0009984306091742447  
Episode: 457, Score: -231.70341335567272, Threshold: 0.0009984306091742447  
Episode: 458, Score: -5.87567806064384, Threshold: 0.0009984306091742447  
Episode: 459, Score: -235.1367232551791, Threshold: 0.0009984306091742447  
Episode: 460, Score: -309.35114854897404, Threshold: 0.0009984306091742447  
Episode: 461, Score: -2.7705127831183054, Threshold: 0.0009984306091742447

```
Episode: 462, Score: -128.42035018761112, Threshold: 0.0009984306091742447
Episode: 463, Score: -123.88832421428859, Threshold: 0.0009984306091742447
Episode: 464, Score: -129.1507713067729, Threshold: 0.0009984306091742447
Episode: 465, Score: -338.59594021445093, Threshold: 0.0009984306091742447
Episode: 466, Score: -130.64855920626934, Threshold: 0.0009984306091742447
Episode: 467, Score: -117.12942985600921, Threshold: 0.0009984306091742447
Episode: 468, Score: -127.8101378770023, Threshold: 0.0009984306091742447
Episode: 469, Score: -127.33773114814598, Threshold: 0.0009984306091742447
Episode: 470, Score: -129.34860117538912, Threshold: 0.0009984306091742447
Episode: 471, Score: -124.07072788366028, Threshold: 0.0009984306091742447
Episode: 472, Score: -228.67090704919005, Threshold: 0.0009984306091742447
Episode: 473, Score: -406.28015688423767, Threshold: 0.0009984306091742447
Episode: 474, Score: -116.98436340497976, Threshold: 0.0009984306091742447
Episode: 475, Score: -348.61906981038106, Threshold: 0.0009984306091742447
Episode: 476, Score: -3.172754327555825, Threshold: 0.0009984306091742447
Episode: 477, Score: -129.083049998916, Threshold: 0.0009984306091742447
Episode: 478, Score: -116.0168383328064, Threshold: 0.0009984306091742447
Episode: 479, Score: -233.36209246143667, Threshold: 0.0009984306091742447
Episode: 480, Score: -4.481225152879583, Threshold: 0.0009984306091742447
Episode: 481, Score: -5.709948734904365, Threshold: 0.0009984306091742447
Episode: 482, Score: -127.83429941421551, Threshold: 0.0009984306091742447
Episode: 483, Score: -128.53677798242822, Threshold: 0.0009984306091742447
Episode: 484, Score: -127.55473502637271, Threshold: 0.0009984306091742447
Episode: 485, Score: -243.4384642651033, Threshold: 0.0009984306091742447
Episode: 486, Score: -122.93125814709515, Threshold: 0.0009984306091742447
Episode: 487, Score: -127.04404331178853, Threshold: 0.0009984306091742447
Episode: 488, Score: -126.06794128560537, Threshold: 0.0009984306091742447
Episode: 489, Score: -125.23594536678952, Threshold: 0.0009984306091742447
Episode: 490, Score: -121.57453819299468, Threshold: 0.0009984306091742447
Episode: 491, Score: -338.93014292828866, Threshold: 0.0009984306091742447
Episode: 492, Score: -4.897988097437928, Threshold: 0.0009984306091742447
Episode: 493, Score: -2.995838769657938, Threshold: 0.0009984306091742447
Episode: 494, Score: -230.94928285646242, Threshold: 0.0009984306091742447
Episode: 495, Score: -131.40632479262743, Threshold: 0.0009984306091742447
Episode: 496, Score: -3.769789911865733, Threshold: 0.0009984306091742447
Episode: 497, Score: -289.50799208843245, Threshold: 0.0009984306091742447
Episode: 498, Score: -236.8091230970299, Threshold: 0.0009984306091742447
Episode: 499, Score: -221.02126227731958, Threshold: 0.0009984306091742447
```



```
In [128]: evaluate_model(finalAgent, 100)
```

Evaluating Episode: 1, Score: -130.8485724129405  
Evaluating Episode: 2, Score: -8.086043676251842  
Evaluating Episode: 3, Score: -132.91841628529534  
Evaluating Episode: 4, Score: -132.20083976079079  
Evaluating Episode: 5, Score: -128.09963226528143  
Evaluating Episode: 6, Score: -231.66940733682839  
Evaluating Episode: 7, Score: -133.37162811611785  
Evaluating Episode: 8, Score: -132.00232955688094  
Evaluating Episode: 9, Score: -231.6252922993193  
Evaluating Episode: 10, Score: -242.14166728843674  
Evaluating Episode: 11, Score: -8.329883901581784  
Evaluating Episode: 12, Score: -127.10209625071896  
Evaluating Episode: 13, Score: -129.31930826405573  
Evaluating Episode: 14, Score: -130.93186152162636  
Evaluating Episode: 15, Score: -121.42012281408948  
Evaluating Episode: 16, Score: -130.70775486074578  
Evaluating Episode: 17, Score: -129.08057334219853  
Evaluating Episode: 18, Score: -8.51425919396035  
Evaluating Episode: 19, Score: -7.912727216808989  
Evaluating Episode: 20, Score: -238.5153605750016  
Evaluating Episode: 21, Score: -254.22943024904586  
Evaluating Episode: 22, Score: -233.96251605130624  
Evaluating Episode: 23, Score: -126.47797443962172  
Evaluating Episode: 24, Score: -231.8426027293189  
Evaluating Episode: 25, Score: -134.80069298316826  
Evaluating Episode: 26, Score: -129.88966242934544  
Evaluating Episode: 27, Score: -122.86602325358098  
Evaluating Episode: 28, Score: -234.55027936179818  
Evaluating Episode: 29, Score: -134.50084694036167  
Evaluating Episode: 30, Score: -125.29161619235558  
Evaluating Episode: 31, Score: -133.60993118770367  
Evaluating Episode: 32, Score: -238.67417686228185  
Evaluating Episode: 33, Score: -8.580048872177867  
Evaluating Episode: 34, Score: -129.7800790119593  
Evaluating Episode: 35, Score: -124.75675512756538  
Evaluating Episode: 36, Score: -126.66294535460698  
Evaluating Episode: 37, Score: -133.2964376286925  
Evaluating Episode: 38, Score: -133.19707482596914  
Evaluating Episode: 39, Score: -131.31048183478194  
Evaluating Episode: 40, Score: -244.07721665022194  
Evaluating Episode: 41, Score: -128.35254747399378  
Evaluating Episode: 42, Score: -122.36707592915333  
Evaluating Episode: 43, Score: -131.64421532594093  
Evaluating Episode: 44, Score: -125.00642269426561  
Evaluating Episode: 45, Score: -128.4765076026033  
Evaluating Episode: 46, Score: -239.06135652743174  
Evaluating Episode: 47, Score: -229.69440706954177  
Evaluating Episode: 48, Score: -240.21242548282243  
Evaluating Episode: 49, Score: -7.916810893939375  
Evaluating Episode: 50, Score: -249.90688735178642  
Evaluating Episode: 51, Score: -129.42646063396356  
Evaluating Episode: 52, Score: -128.88928400396753  
Evaluating Episode: 53, Score: -1841.8681697985533  
Evaluating Episode: 54, Score: -121.88322476917797  
Evaluating Episode: 55, Score: -130.31642308839696  
Evaluating Episode: 56, Score: -248.41921368753907  
Evaluating Episode: 57, Score: -123.07527838666242  
Evaluating Episode: 58, Score: -131.68741339485288  
Evaluating Episode: 59, Score: -132.96196078413405  
Evaluating Episode: 60, Score: -133.7352696006326  
Evaluating Episode: 61, Score: -8.524410696437721  
Evaluating Episode: 62, Score: -132.68196851080845  
Evaluating Episode: 63, Score: -129.69709731142777  
Evaluating Episode: 64, Score: -132.97824676083  
Evaluating Episode: 65, Score: -132.49316072751864  
Evaluating Episode: 66, Score: -123.47050275383542

```
Evaluating Episode: 67, Score: -131.57786406470325
Evaluating Episode: 68, Score: -246.8130920038204
Evaluating Episode: 69, Score: -132.6583708119997
Evaluating Episode: 70, Score: -238.86658917671693
Evaluating Episode: 71, Score: -134.3770728821542
Evaluating Episode: 72, Score: -7.942608648246738
Evaluating Episode: 73, Score: -8.091996170688677
Evaluating Episode: 74, Score: -132.57004023431614
Evaluating Episode: 75, Score: -135.3195799602425
Evaluating Episode: 76, Score: -129.30082342999629
Evaluating Episode: 77, Score: -8.036891696520437
Evaluating Episode: 78, Score: -128.85376766760578
Evaluating Episode: 79, Score: -133.59369463913808
Evaluating Episode: 80, Score: -131.36068978856315
Evaluating Episode: 81, Score: -129.06857961641037
Evaluating Episode: 82, Score: -126.43382639988315
Evaluating Episode: 83, Score: -125.69561760069196
Evaluating Episode: 84, Score: -119.84318626264454
Evaluating Episode: 85, Score: -131.7782591841452
Evaluating Episode: 86, Score: -251.90639575758183
Evaluating Episode: 87, Score: -132.10121313023345
Evaluating Episode: 88, Score: -125.73492420007791
Evaluating Episode: 89, Score: -233.4007877893496
Evaluating Episode: 90, Score: -122.19944634310747
Evaluating Episode: 91, Score: -127.47998539835308
Evaluating Episode: 92, Score: -126.95203254299423
Evaluating Episode: 93, Score: -126.95260319151612
Evaluating Episode: 94, Score: -8.195556496887534
Evaluating Episode: 95, Score: -7.915958621786024
Evaluating Episode: 96, Score: -1850.0886253806118
Evaluating Episode: 97, Score: -9.975444311075977
Evaluating Episode: 98, Score: -238.5089947878025
Evaluating Episode: 99, Score: -131.17847069903485
Evaluating Episode: 100, Score: -124.77295283743057
Final Score: -170.01447251911344
```

```
In [129...]: visualize_model(finalAgent, 3)
```

```
Visualizing Episode: 1, Score: -126.29451481752481
Visualizing Episode: 2, Score: -246.16612591783476
Visualizing Episode: 3, Score: -270.686294221618
```

## Conclusion

In conclusion, we have learnt how to use DQN to create a Reinforcement Learning algorithm that can balance a pendulum. At the start it was really frustrating as the simple DQN kept giving inconsistent results. But as we applied improvements both to stabilisation and architecture, we saw substantial improvement in the model.

An essential takeaway from this experience is the understanding that the simplicity of the pendulum problem limits the effectiveness of advanced methods like dueling and prioritized experience replay (PER). In cases where the environment itself is straightforward, the additional complexity introduced by these methods may not yield substantial benefits. Thus, the choice of reinforcement learning techniques should align with the complexity of the problem at hand to strike an optimal balance between simplicity and effectiveness.