ate-for-actor-critic-problem-final

September 19, 2024

0.0.1 Group ID: 71

0.0.2 Group Members Name with Student ID:

(1) S IYNESHDURAI - 2023AA05192 - 100% Contribution

(2) AKBARSHA M - 2023AA05159 - 100% Contribution

(3) MITHUN S - 2023AB05188 - 100% Contribution

(4) ARUL PANDIYAN K - 2023AA05146 - 100% Contribution

1 Problem Statement

The objective of the problem is to implement an Actor-Critic reinforcement learning algorithm to optimize energy consumption in a building. The agent should learn to adjust the temperature settings dynamically to minimize energy usage while maintaining comfortable indoor conditions.

Dataset Details Dataset: https://archive.ics.uci.edu/dataset/374/appliances+energy+prediction

This dataset contains energy consumption data for a residential building, along with various environmental and operational factors.

Data Dictionary: * Appliances: Energy use in Wh * lights: Energy use of light fixtures in the house in Wh * T1 - T9: Temperatures in various rooms and outside * RH_1 to RH_9: Humidity measurements in various rooms and outside * Visibility: Visibility in km * Tdewpoint: Dew point temperature * Pressure_mm_hgg: Pressure in mm Hg * Windspeed: Wind speed in m/s

Environment Details **State Space:** The state space consists of various features from the dataset that impact energy consumption and comfort levels.

- Current Temperature (T1 to T9): Temperatures in various rooms and outside.
- Current Humidity (RH_1 to RH_9): Humidity measurements in different locations.
- Visibility (Visibility): Visibility in meters.
- Dew Point (Tdewpoint): Dew point temperature.
- Pressure (Press_mm_hg): Atmospheric pressure in mm Hg.
- Windspeed (Windspeed): Wind speed in m/s.

Total State Vector Dimension: Number of features = 9 (temperature) + 9 (humidity) + 1 (visibility) + 1 (dew point) + 1 (pressure) + 1 (windspeed) = 22 features

Target Variable: Appliances (energy consumption in Wh).

Action Space: The action space consists of discrete temperature adjustments: * Action 0: Decrease temperature by 1°C * Action 1: Maintain current temperature * Action 2: Increase temperature by 1°C

- If the action is to decrease the temperature by 1°C, you'll adjust each temperature feature (T1 to T9) down by 1°C.
- If the action is to increase the temperature by 1°C, you'll adjust each temperature feature (T1 to T9) up by 1°C.
- Other features remain unchanged.

Policy (Actor): A neural network that outputs a probability distribution over possible temperature adjustments.

Value function (Critic): A neural network that estimates the expected cumulative reward (energy savings) from a given state.

Reward function: The reward function should reflect the overall comfort and energy efficiency based on all temperature readings. i.e., balance between minimising temperature deviations and minimizing energy consumption.

- Calculate the penalty based on the deviation of each temperature from the target temperature and then aggregate these penalties.
- Measure the change in energy consumption before and after applying the RL action.
- Combine the comfort penalty and energy savings to get the final reward.

Example:

Target temperature=22°C

Initial Temperatures: T1=23, T2=22, T3=21, T4=23, T5=22, T6=21, T7=24, T8=22, T9=23

Action Taken: Decrease temperature by 1°C for each room

Resulting Temperatures: T1 = 22, T2 = 21, T3 = 20, T4 = 22, T5 = 21, T6 = 20, T7 = 23, T8 = 21, T9 = 22

Energy Consumption: 50 Wh (before RL adjustment) and 48 Wh (after RL adjustment) * Energy Before (50 Wh): Use the energy consumption from the dataset at the current time step. * Energy After (48 Wh): Use the energy consumption from the dataset at the next time step (if available).

Consider only temperature features for deviation calculation.

Deviation = abs (Ti− Ttarget)

Deviations=[abs(22−22), abs(21−22), abs(20−22), abs(22−22), abs(21−22), abs(20−22), abs(23−22), abs(21−22), abs(22−22)]

Deviations = [0, 1, 2, 0, 1, 2, 1, 1, 0], Sum of deviations = 8

Energy Savings = Energy Before−Energy After = 50 – 48 = 2Wh

Reward= −Sum of Deviations + Energy Savings = -8+6 = -2

Expected Outcomes

1. Pre-process the dataset to handle any missing values and create training and testing sets.

2. Implement the Actor-Critic algorithm using TensorFlow.
3. Train the model over 500 episodes to minimize energy consumption while maintaining an indoor temperature of 22°C.
4. Plot the total reward obtained in each episode to evaluate the learning progress.
5. Evaluate the performance of the model on test set to measure its performance
6. Provide graphs showing the convergence of the Actor and Critic losses.
7. Plot the learned policy by showing the action probabilities across different state values (e.g., temperature settings).
8. Provide an analysis on a comparison of the energy consumption before and after applying the reinforcement learning algorithm.

Code Execution

[ ]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import numpy as np

# Load the dataset
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/00374/
 ↪energydata_complete.csv" # updated URL
data = pd.read_csv(url)

# Fill missing values, if any
data.fillna(method='ffill', inplace=True)

# Define feature groups
temp_features = ['T1', 'T2', 'T3', 'T4', 'T5', 'T6', 'T7', 'T8', 'T9']
other_features = ['RH_1', 'RH_2', 'RH_3', 'RH_4', 'RH_5', 'RH_6', 'RH_7',␣
 ↪'RH_8', 'RH_9',
                  'Visibility', 'Tdewpoint', 'Press_mm_hg', 'Windspeed']
target = ['Appliances']

# Extract features and target
X_temp = data[temp_features].values  # Shape: (n_samples, 9)
X_other = data[other_features].values  # Shape: (n_samples, 13)
y = data[target].values  # Shape: (n_samples, 1)

# Initialize scalers
scaler_other = StandardScaler()

# Scale only the non-temperature features
X_other_scaled = scaler_other.fit_transform(X_other)

# Concatenate temperature (unscaled) and other (scaled) features
X = np.hstack((X_temp, X_other_scaled))  # Shape: (n_samples, 22)

# No scaling for y to maintain energy consumption in original units
```

```
# If you prefer scaling y, ensure to inverse transform when calculating rewards
# For this implementation, we'll keep y unscaled

# Split the data into training and testing sets (80% training, 20% testing)
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=0
)

print(f"Training data shape: {X_train.shape}")
print(f"Testing data shape: {X_test.shape}")
```

```
Training data shape: (15788, 22)
Testing data shape: (3947, 22)
```

```
<ipython-input-9-3df30f27d945>:11: FutureWarning: DataFrame.fillna with 'method'
is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill()
instead.
  data.fillna(method='ffill', inplace=True)
```

Defining Actor Critic Model using tensorflow (1 M)

[ ]:
```
import tensorflow as tf
from tensorflow.keras import layers, models, optimizers

# Define the Energy Prediction Model
def build_energy_model():
    model = models.Sequential([
        layers.Dense(128, activation='relu', input_shape=(22,)),
        layers.Dense(128, activation='relu'),
        layers.Dense(1, activation='linear')  # Predicts energy consumption
    ])
    model.compile(optimizer=optimizers.Adam(learning_rate=0.001), loss='mse')
    return model

# Instantiate and train the energy model
energy_model = build_energy_model()
energy_model.summary()

# Train the energy model
energy_model.fit(
    X_train, y_train,
    epochs=50,  # Increase epochs for better training
    batch_size=64,
    validation_split=0.2,
    verbose=1
)
```

```
/usr/local/lib/python3.10/dist-packages/keras/src/layers/core/dense.py:87:
UserWarning: Do not pass an `input_shape`/`input_dim` argument to a layer. When
```

```
using Sequential models, prefer using an `Input(shape)` object as the first
layer in the model instead.
  super().__init__(activity_regularizer=activity_regularizer, **kwargs)
```

Model: "sequential"

```
 Layer (type)                          Output Shape                          ␣
↪Param #

 dense (Dense)                         (None, 128)                             ␣
↪2,944

 dense_1 (Dense)                       (None, 128)                            ␣
↪16,512

 dense_2 (Dense)                       (None, 1)                                 ␣
↪129
```

Total params: 19,585 (76.50 KB)

Trainable params: 19,585 (76.50 KB)

Non-trainable params: 0 (0.00 B)

```
Epoch 1/50
198/198              4s 7ms/step -
loss: 12273.9570 - val_loss: 10587.6348
Epoch 2/50
198/198              0s 2ms/step -
loss: 9780.0859 - val_loss: 10387.8682
Epoch 3/50
198/198              1s 2ms/step -
loss: 9863.1133 - val_loss: 10066.8262
Epoch 4/50
198/198              1s 2ms/step -
loss: 9468.0049 - val_loss: 9780.3916
Epoch 5/50
198/198              1s 2ms/step -
loss: 9588.0049 - val_loss: 9728.1543
Epoch 6/50
198/198              1s 2ms/step -
loss: 8812.9014 - val_loss: 9651.2461
Epoch 7/50
```

```
198/198             1s 2ms/step -
loss: 9167.1377 - val_loss: 9523.4404
Epoch 8/50
198/198             1s 2ms/step -
loss: 9153.8418 - val_loss: 9420.3154
Epoch 9/50
198/198             1s 2ms/step -
loss: 9573.5596 - val_loss: 9500.3369
Epoch 10/50
198/198             0s 2ms/step -
loss: 9007.2422 - val_loss: 9335.1748
Epoch 11/50
198/198             0s 2ms/step -
loss: 9028.4512 - val_loss: 9364.0996
Epoch 12/50
198/198             1s 2ms/step -
loss: 8543.2988 - val_loss: 9267.5762
Epoch 13/50
198/198             0s 2ms/step -
loss: 9083.5469 - val_loss: 9255.2148
Epoch 14/50
198/198             1s 2ms/step -
loss: 9387.7451 - val_loss: 9293.2295
Epoch 15/50
198/198             0s 2ms/step -
loss: 8364.8184 - val_loss: 9139.6064
Epoch 16/50
198/198             1s 2ms/step -
loss: 9363.8604 - val_loss: 9139.3447
Epoch 17/50
198/198             1s 3ms/step -
loss: 8769.6758 - val_loss: 9097.0498
Epoch 18/50
198/198             1s 3ms/step -
loss: 8716.0322 - val_loss: 9104.8330
Epoch 19/50
198/198             1s 3ms/step -
loss: 8844.4541 - val_loss: 9109.4014
Epoch 20/50
198/198             1s 3ms/step -
loss: 8476.3662 - val_loss: 9008.9561
Epoch 21/50
198/198             1s 2ms/step -
loss: 8767.2910 - val_loss: 9017.6318
Epoch 22/50
198/198             1s 2ms/step -
loss: 8564.5010 - val_loss: 8917.7441
Epoch 23/50
```

```
198/198             1s 2ms/step -
loss: 9217.2793 - val_loss: 8954.7158
Epoch 24/50
198/198             0s 2ms/step -
loss: 8791.2754 - val_loss: 9029.3447
Epoch 25/50
198/198             1s 2ms/step -
loss: 8655.0742 - val_loss: 8854.3047
Epoch 26/50
198/198             0s 2ms/step -
loss: 8251.3525 - val_loss: 9152.3057
Epoch 27/50
198/198             1s 2ms/step -
loss: 8729.6016 - val_loss: 8798.8242
Epoch 28/50
198/198             1s 2ms/step -
loss: 8047.7207 - val_loss: 8968.5107
Epoch 29/50
198/198             0s 2ms/step -
loss: 8631.7324 - val_loss: 8835.3086
Epoch 30/50
198/198             0s 2ms/step -
loss: 9042.0439 - val_loss: 8806.4141
Epoch 31/50
198/198             0s 2ms/step -
loss: 8299.9209 - val_loss: 8869.6699
Epoch 32/50
198/198             0s 2ms/step -
loss: 8344.1738 - val_loss: 9005.8086
Epoch 33/50
198/198             1s 4ms/step -
loss: 8484.0010 - val_loss: 8619.5098
Epoch 34/50
198/198             0s 2ms/step -
loss: 7995.5435 - val_loss: 8653.8330
Epoch 35/50
198/198             1s 3ms/step -
loss: 7938.4727 - val_loss: 8710.1943
Epoch 36/50
198/198             0s 2ms/step -
loss: 8894.8350 - val_loss: 8592.9502
Epoch 37/50
198/198             0s 2ms/step -
loss: 7947.5498 - val_loss: 8666.6641
Epoch 38/50
198/198             1s 2ms/step -
loss: 8206.3994 - val_loss: 9523.1738
Epoch 39/50
```

```
198/198              1s 3ms/step -
loss: 8584.0918 - val_loss: 8588.3447
Epoch 40/50
198/198              1s 2ms/step -
loss: 8275.7920 - val_loss: 8559.8613
Epoch 41/50
198/198              1s 3ms/step -
loss: 8419.5713 - val_loss: 8532.5732
Epoch 42/50
198/198              1s 3ms/step -
loss: 7894.7114 - val_loss: 8557.6133
Epoch 43/50
198/198              0s 2ms/step -
loss: 7474.5991 - val_loss: 8465.7334
Epoch 44/50
198/198              1s 2ms/step -
loss: 8241.5195 - val_loss: 8490.3857
Epoch 45/50
198/198              1s 2ms/step -
loss: 8276.1406 - val_loss: 8400.1865
Epoch 46/50
198/198              0s 2ms/step -
loss: 7715.3735 - val_loss: 8422.7812
Epoch 47/50
198/198              1s 2ms/step -
loss: 7872.0059 - val_loss: 8486.9346
Epoch 48/50
198/198              1s 2ms/step -
loss: 8236.2842 - val_loss: 8672.3398
Epoch 49/50
198/198              0s 2ms/step -
loss: 8441.2393 - val_loss: 8333.0820
Epoch 50/50
198/198              0s 2ms/step -
loss: 8291.7129 - val_loss: 8414.0371
```

[ ]: <keras.src.callbacks.history.History at 0x7b6e16f008b0>

1.0.1 Reward Function (0.5 M)

[ ]:
```
def calculate_reward(state, next_state, y_current, y_next, target_temp=22):
    """
    Calculate the reward based on temperature deviations and energy savings.

    Parameters:
    - state: Current state vector (numpy array)
    - next_state: Next state vector after action (numpy array)
```

```
    - y_current: Current energy consumption (float)
    - y_next: Next energy consumption after action (float)
    - target_temp: Desired target temperature (float)

    Returns:
    - reward: Calculated reward (float)
    """
    # Calculate sum of absolute deviations for all temperature features
    temp_deviation = np.sum(np.abs(next_state[:9] - target_temp))

    # Calculate energy savings
    energy_saving = y_current - y_next  # Positive if energy decreased

    # Define reward as negative deviations plus energy savings
    reward = -temp_deviation + energy_saving

    return reward
```

Environment Simulation (0.5 M)

[ ]:
```
def simulate_environment(state, action, energy_model, target_temp=22):
    """
    Simulate the environment's response to an action.

    Parameters:
    - state: Current state vector (numpy array)
    - action: Action taken (int: 0=Decrease, 1=Maintain, 2=Increase)
    - energy_model: Trained energy prediction model
    - target_temp: Desired target temperature (float)

    Returns:
    - next_state: New state vector after action (numpy array)
    - y_next: Predicted energy consumption after action (float)
    - reward: Calculated reward (float)
    """
    # Define temperature adjustments based on action
    temp_adjustment = [-1, 0, 1]  # Decrease, Maintain, Increase

    # Create a copy of the current state to modify
    next_state = state.copy()

    # Apply temperature adjustment to all temperature features (T1 to T9)
    next_state[:9] += temp_adjustment[action]

    # Predict energy consumption for the next state
    y_next = energy_model.predict(next_state.reshape(1, -1))[0][0]
```

```python
    # Predict current energy consumption
    y_current = energy_model.predict(state.reshape(1, -1))[0][0]

    # Calculate reward
    reward = calculate_reward(state, next_state, y_current, y_next,␣
↪target_temp=target_temp)

    return next_state, y_next, reward
```

[ ]:
```python
from tensorflow.keras import layers, models, optimizers

# Define the Actor model
def build_actor_model(state_space, action_space):
    model = models.Sequential([
        layers.InputLayer(input_shape=(state_space,)),
        layers.Dense(128, activation='relu'),
        layers.Dense(128, activation='relu'),
        layers.Dense(action_space, activation='softmax')  # Outputs action␣
↪probabilities
    ])
    model.compile(optimizer=optimizers.Adam(learning_rate=0.001))
    return model
```

[ ]:
```python
# Define the Critic model
def build_critic_model(state_space):
    model = models.Sequential([
        layers.InputLayer(input_shape=(state_space,)),
        layers.Dense(128, activation='relu'),
        layers.Dense(128, activation='relu'),
        layers.Dense(1, activation='linear')  # Outputs state-value
    ])
    model.compile(optimizer=optimizers.Adam(learning_rate=0.001), loss='mse')
    return model
```

[ ]:
```python
# Define state and action spaces
state_space = X_train.shape[1]  # 22 features
action_space = 3  # 0=Decrease, 1=Maintain, 2=Increase

# Instantiate Actor and Critic models
actor_model = build_actor_model(state_space, action_space)
critic_model = build_critic_model(state_space)

# Display model summaries
actor_model.summary()
critic_model.summary()
```

```
/usr/local/lib/python3.10/dist-packages/keras/src/layers/core/input_layer.py:26:
```

```
UserWarning: Argument `input_shape` is deprecated. Use `shape` instead.
  warnings.warn(
```

```
Model: "sequential_1"


 Layer (type)                         Output Shape                          ␣
↪Param #

 dense_3 (Dense)                      (None, 128)                             ␣
↪2,944

 dense_4 (Dense)                      (None, 128)                            ␣
↪16,512

 dense_5 (Dense)                      (None, 3)                                 ␣
↪387


 Total params: 19,843 (77.51 KB)

 Trainable params: 19,843 (77.51 KB)

 Non-trainable params: 0 (0.00 B)

Model: "sequential_2"


 Layer (type)                         Output Shape                          ␣
↪Param #

 dense_6 (Dense)                      (None, 128)                             ␣
↪2,944

 dense_7 (Dense)                      (None, 128)                            ␣
↪16,512

 dense_8 (Dense)                      (None, 1)                                 ␣
↪129


 Total params: 19,585 (76.50 KB)
```

Trainable params: 19,585 (76.50 KB)

Non-trainable params: 0 (0.00 B)

Implementation of Training Function (2 M)

[ ]:
```
import tensorflow as tf
import numpy as np

def train_actor_critic(actor_model, critic_model, energy_model, X_train,␣
↪episodes=500, gamma=0.99):
    actor_losses = []
    critic_losses = []

    # Compile critic model with an optimizer
    critic_model.compile(optimizer=tf.keras.optimizers.Adam())

    for episode in range(episodes):
        total_reward = 0
        state = X_train[np.random.choice(len(X_train))]  # Random initial state

        with tf.GradientTape(persistent=True) as tape_critic, tf.GradientTape()␣
↪as tape_actor:
            # Predict critic value
            critic_value = critic_model(state.reshape(1, -1))

            # Simulate an action (e.g., random action for now)
            action = np.random.choice([0, 1])  # Assuming binary action space

            # Get target value based on reward from the energy model and next␣
↪state
            # Use forward pass through energy model instead of predict()
            reward = energy_model(state.reshape(1, -1), training=True)  #␣
↪Forward pass to maintain gradient flow
            next_state = X_train[np.random.choice(len(X_train))]  # Simulate␣
↪the next state
            next_value = critic_model(next_state.reshape(1, -1), training=True)␣
↪ # Forward pass to maintain gradients

            # Compute target value: reward + gamma * next_value
            target_value = reward + gamma * next_value
            total_reward += reward

            # Compute critic loss
            # Use tf.keras.losses.MeanSquaredError() to create a loss object
```

```
            critic_loss = tf.keras.losses.MeanSquaredError()(target_value,␣
↪critic_value)

            # Compute actor loss (using advantage)
            advantage = target_value - critic_value
            log_probs = actor_model(state.reshape(1, -1), training=True)  #␣
↪Forward pass to maintain gradients
            actor_loss = -log_probs * advantage

        # Compute and apply gradients for the critic
        grads_critic = tape_critic.gradient(critic_loss, critic_model.
↪trainable_variables)
        critic_model.optimizer.apply_gradients(zip(grads_critic, critic_model.
↪trainable_variables))

        # Compute and apply gradients for the actora
        actor_optimizer = tf.keras.optimizers.Adam()  # Define an optimizer for␣
↪actor
        grads_actor = tape_actor.gradient(actor_loss, actor_model.
↪trainable_variables)
        actor_optimizer.apply_gradients(zip(grads_actor, actor_model.
↪trainable_variables))

        # Store losses
        actor_losses.append(actor_loss.numpy())
        critic_losses.append(critic_loss.numpy())

        # Logging each episode's details
        print(f"Episode {episode+1}/{episodes}, Total Reward: {total_reward},␣
↪Actor Loss: {actor_loss.numpy()}, Critic Loss: {critic_loss.numpy()}")

    return actor_losses, critic_losses
```

[ ]:
```
# Call training function
actor_losses, critic_losses = train_actor_critic(actor_model, critic_model,␣
 ↪energy_model, X_train, episodes=500, gamma=0.99)
```

```
Episode 1/500, Total Reward: [[91.405]], Actor Loss: [[ -0.12714262 -88.07152
-1.7712694 ]], Critic Loss: 8094.58740234375
Episode 2/500, Total Reward: [[98.495155]], Actor Loss: [[ -6.314211 -43.78624
-49.293377]], Critic Loss: 9879.1318359375
Episode 3/500, Total Reward: [[185.08284]], Actor Loss: [[ -13.121733 -152.68196
-18.284948]], Critic Loss: 33888.625
Episode 4/500, Total Reward: [[107.9656]], Actor Loss: [[-21.557621 -61.24501
-28.057219]], Critic Loss: 12289.90625
Episode 5/500, Total Reward: [[48.528103]], Actor Loss: [[-13.425478  -22.47305
-13.0330925]], Critic Loss: 2394.30322265625
```

```
Episode 6/500, Total Reward: [[113.69542]], Actor Loss: [[-24.671078 -67.29334
-19.136793]], Critic Loss: 12343.4794921875
Episode 7/500, Total Reward: [[77.32563]], Actor Loss: [[-19.690214 -32.828503
-24.253145]], Critic Loss: 5893.91845703125
Episode 8/500, Total Reward: [[77.572685]], Actor Loss: [[-31.885958 -20.765062
-27.2844  ]], Critic Loss: 6389.6708984375
Episode 9/500, Total Reward: [[43.888752]], Actor Loss: [[-13.156229 -12.716449
-17.500517]], Critic Loss: 1881.2337646484375
Episode 10/500, Total Reward: [[32.421963]], Actor Loss: [[-12.169235  -7.762091
-13.805387]], Critic Loss: 1138.165771484375
Episode 11/500, Total Reward: [[201.019]], Actor Loss: [[-110.446434  -39.892056
-51.40605 ]], Critic Loss: 40700.859375
Episode 12/500, Total Reward: [[63.926395]], Actor Loss: [[-22.28458  -30.891441
-11.547242]], Critic Loss: 4189.10107421875
Episode 13/500, Total Reward: [[102.81784]], Actor Loss: [[-19.363407 -54.803505
-26.220875]], Critic Loss: 10077.7080078125
Episode 14/500, Total Reward: [[82.49328]], Actor Loss: [[-31.433504 -15.913264
-36.695553]], Critic Loss: 7063.11181640625
Episode 15/500, Total Reward: [[82.210014]], Actor Loss: [[-16.970818  -50.0057
-7.1389937]], Critic Loss: 5493.10888671875
Episode 16/500, Total Reward: [[88.793526]], Actor Loss: [[-24.998198 -42.00946
-23.085371]], Critic Loss: 8116.75341796875
Episode 17/500, Total Reward: [[76.68712]], Actor Loss: [[-50.365196  -9.512346
-17.972237]], Critic Loss: 6060.587890625
Episode 18/500, Total Reward: [[44.749485]], Actor Loss: [[-11.4320755
-15.891979  -12.385938 ]], Critic Loss: 1576.8834228515625
Episode 19/500, Total Reward: [[62.54498]], Actor Loss: [[-36.689392  -8.366429
-16.728518]], Critic Loss: 3817.3046875
Episode 20/500, Total Reward: [[75.311485]], Actor Loss: [[-47.25539
-7.0786786 -20.210272 ]], Critic Loss: 5556.85888671875
Episode 21/500, Total Reward: [[54.717304]], Actor Loss: [[-33.32344   -7.412706
-14.427091]], Critic Loss: 3042.98291015625
Episode 22/500, Total Reward: [[93.02986]], Actor Loss: [[-68.871956   -8.931289
-7.3363986]], Critic Loss: 7248.759765625
Episode 23/500, Total Reward: [[74.79785]], Actor Loss: [[-71.44517
-0.6976478  -2.2873046]], Critic Loss: 5539.84326171875
Episode 24/500, Total Reward: [[100.92501]], Actor Loss: [[-97.07623
-1.8011203  -8.308875 ]], Critic Loss: 11488.8857421875
Episode 25/500, Total Reward: [[283.1216]], Actor Loss: [[-215.54971   -31.29688
-27.749125]], Critic Loss: 75402.796875
Episode 26/500, Total Reward: [[88.15415]], Actor Loss: [[-60.341923  -7.72378
-17.690493]], Critic Loss: 7354.125
Episode 27/500, Total Reward: [[41.819958]], Actor Loss: [[-25.243067
-3.9488719 -11.889284 ]], Critic Loss: 1687.6668701171875
Episode 28/500, Total Reward: [[85.77897]], Actor Loss: [[-70.05963   -4.227087
-10.653404]], Critic Loss: 7214.82470703125
Episode 29/500, Total Reward: [[82.631386]], Actor Loss: [[-52.336613 -10.587441
-20.542814]], Critic Loss: 6966.7177734375
```

```
Episode 30/500, Total Reward: [[104.785774]], Actor Loss: [[-82.97617
-4.7690096 -14.738164 ]], Critic Loss: 10502.8359375
Episode 31/500, Total Reward: [[129.71303]], Actor Loss: [[-31.70655  -31.563183
-71.76883 ]], Critic Loss: 18235.412109375
Episode 32/500, Total Reward: [[74.089676]], Actor Loss: [[-15.698352 -23.327698
-35.897846]], Critic Loss: 5613.59033203125
Episode 33/500, Total Reward: [[160.46294]], Actor Loss: [[ -23.850096
-35.47195  -100.705864]], Critic Loss: 25608.931640625
Episode 34/500, Total Reward: [[199.37129]], Actor Loss: [[ -34.25863
-55.764618 -108.59805 ]], Critic Loss: 39450.41796875
Episode 35/500, Total Reward: [[74.81722]], Actor Loss: [[-55.57836   -9.058722
-8.63702 ]], Critic Loss: 5369.09375
Episode 36/500, Total Reward: [[75.69022]], Actor Loss: [[ -1.3093767 -59.533413
-3.0981965]], Critic Loss: 4088.44970703125
Episode 37/500, Total Reward: [[99.38407]], Actor Loss: [[ -9.466052 -75.45498
-4.909153]], Critic Loss: 8069.4619140625
Episode 38/500, Total Reward: [[281.11633]], Actor Loss: [[ -48.48533  -147.2035
-79.479805]], Critic Loss: 75717.78125
Episode 39/500, Total Reward: [[100.24818]], Actor Loss: [[-13.330101 -60.14442
-42.3926  ]], Critic Loss: 13425.189453125
Episode 40/500, Total Reward: [[148.87262]], Actor Loss: [[ -15.445616
-105.20261   -25.349743]], Critic Loss: 21315.408203125
Episode 41/500, Total Reward: [[58.20014]], Actor Loss: [[-14.204674 -26.464329
-18.30593 ]], Critic Loss: 3478.042724609375
Episode 42/500, Total Reward: [[104.052666]], Actor Loss: [[ -8.367683 -79.70144
-14.467784]], Critic Loss: 10513.818359375
Episode 43/500, Total Reward: [[82.454384]], Actor Loss: [[ -1.6487176 -80.08987
-9.555674 ]], Critic Loss: 8334.642578125
Episode 44/500, Total Reward: [[70.546486]], Actor Loss: [[-2.2304529e-02
-6.5508736e+01 -2.4277003e-01]], Critic Loss: 4326.193359375
Episode 45/500, Total Reward: [[83.069756]], Actor Loss: [[-5.6343116e-02
-8.2978493e+01 -2.8127763e-01]], Critic Loss: 6941.57373046875
Episode 46/500, Total Reward: [[66.09931]], Actor Loss: [[ -0.35282218 -50.05212
-2.370658  ]], Critic Loss: 2785.263671875
Episode 47/500, Total Reward: [[120.174324]], Actor Loss: [[-12.3919525
-67.72245   -27.41971  ]], Critic Loss: 11563.5869140625
Episode 48/500, Total Reward: [[144.84175]], Actor Loss: [[ -7.879466 -87.062675
-44.433628]], Critic Loss: 19425.603515625
Episode 49/500, Total Reward: [[89.01855]], Actor Loss: [[ -3.433242 -70.417984
-15.10868 ]], Critic Loss: 7913.865234375
Episode 50/500, Total Reward: [[57.45492]], Actor Loss: [[ -0.62799716 -48.28838
-9.488295  ]], Critic Loss: 3411.10546875
Episode 51/500, Total Reward: [[129.11848]], Actor Loss: [[  -2.0325828
-107.29313    -19.787477 ]], Critic Loss: 16670.21484375
Episode 52/500, Total Reward: [[71.42916]], Actor Loss: [[ -2.282933 -46.902866
-16.904778]], Critic Loss: 4367.96435546875
Episode 53/500, Total Reward: [[59.307175]], Actor Loss: [[ -3.231874 -44.232113
-16.594162]], Critic Loss: 4103.44580078125
```

```
Episode 54/500, Total Reward: [[141.33675]], Actor Loss: [[-19.382954 -65.236084
-76.22715 ]], Critic Loss: 25871.498046875
Episode 55/500, Total Reward: [[107.451965]], Actor Loss: [[-11.833614 -70.53321
-25.124113]], Critic Loss: 11554.3017578125
Episode 56/500, Total Reward: [[54.222958]], Actor Loss: [[ -6.225319 -19.61307
-23.29845 ]], Critic Loss: 2414.4287109375
Episode 57/500, Total Reward: [[84.370926]], Actor Loss: [[-20.491777 -49.563347
-19.694466]], Critic Loss: 8054.98876953125
Episode 58/500, Total Reward: [[176.39877]], Actor Loss: [[-58.479378 -61.828117
-56.549442]], Critic Loss: 31278.375
Episode 59/500, Total Reward: [[52.06536]], Actor Loss: [[-20.328373 -26.750015
-23.971268]], Critic Loss: 5048.05322265625
Episode 60/500, Total Reward: [[62.20805]], Actor Loss: [[-19.574324 -22.012463
-20.27882 ]], Critic Loss: 3827.35302734375
Episode 61/500, Total Reward: [[54.86623]], Actor Loss: [[-20.916212 -15.05451
-12.156869]], Critic Loss: 2316.265380859375
Episode 62/500, Total Reward: [[79.00108]], Actor Loss: [[-18.269043 -39.130436
-19.189985]], Critic Loss: 5865.94580078125
Episode 63/500, Total Reward: [[189.52837]], Actor Loss: [[-54.27654  -96.925934
-38.115856]], Critic Loss: 35841.4296875
Episode 64/500, Total Reward: [[113.17004]], Actor Loss: [[-22.336533 -70.783394
-14.843282]], Critic Loss: 11656.0546875
Episode 65/500, Total Reward: [[103.76655]], Actor Loss: [[-29.154406 -49.045227
-24.16031 ]], Critic Loss: 10477.55859375
Episode 66/500, Total Reward: [[36.046177]], Actor Loss: [[ -9.3299885
-20.382309   -6.0653667]], Critic Loss: 1280.04150390625
Episode 67/500, Total Reward: [[91.95689]], Actor Loss: [[-14.888192 -70.08174
-7.843543]], Critic Loss: 8614.341796875
Episode 68/500, Total Reward: [[91.366554]], Actor Loss: [[ -2.27574
-85.479416   -1.2748839]], Critic Loss: 7926.34765625
Episode 69/500, Total Reward: [[181.1928]], Actor Loss: [[-78.88133  -51.514454
-37.85022 ]], Critic Loss: 28306.716796875
Episode 70/500, Total Reward: [[64.30413]], Actor Loss: [[-22.992754 -26.303358
-20.496527]], Critic Loss: 4871.0126953125
Episode 71/500, Total Reward: [[77.38914]], Actor Loss: [[-29.58815  -26.928438
-22.5676  ]], Critic Loss: 6254.30810546875
Episode 72/500, Total Reward: [[60.2665]], Actor Loss: [[-26.339222 -15.467176
-13.731636]], Critic Loss: 3084.47314453125
Episode 73/500, Total Reward: [[59.038403]], Actor Loss: [[-20.69068  -26.344538
-20.576706]], Critic Loss: 4571.3720703125
Episode 74/500, Total Reward: [[59.34159]], Actor Loss: [[-25.646639 -24.418388
-24.654596]], Critic Loss: 5583.021484375
Episode 75/500, Total Reward: [[125.90421]], Actor Loss: [[-52.432358 -30.919563
-48.849567]], Critic Loss: 17477.234375
Episode 76/500, Total Reward: [[34.55101]], Actor Loss: [[-17.935177 -14.400395
-21.575901]], Critic Loss: 2906.44677734375
Episode 77/500, Total Reward: [[63.286213]], Actor Loss: [[-23.670675 -22.033205
-21.898018]], Critic Loss: 4570.0166015625
```

```
Episode 78/500, Total Reward: [[66.90808]], Actor Loss: [[-24.213312 -20.668133
-13.422341]], Critic Loss: 3399.3310546875
Episode 79/500, Total Reward: [[157.33656]], Actor Loss: [[-52.90057 -56.87295
-38.00178]], Critic Loss: 21837.5390625
Episode 80/500, Total Reward: [[66.10409]], Actor Loss: [[-23.744822 -20.217676
-26.500795]], Critic Loss: 4965.076171875
Episode 81/500, Total Reward: [[49.902157]], Actor Loss: [[-23.035933 -14.870735
-21.400076]], Critic Loss: 3517.289794921875
Episode 82/500, Total Reward: [[74.483665]], Actor Loss: [[-36.581554 -12.10802
-21.036322]], Critic Loss: 4861.701171875
Episode 83/500, Total Reward: [[102.33906]], Actor Loss: [[-25.361979 -47.535805
-4.67074 ]], Critic Loss: 6016.87646484375
Episode 84/500, Total Reward: [[72.68705]], Actor Loss: [[-23.513714 -43.966515
-14.321691]], Critic Loss: 6691.5537109375
Episode 85/500, Total Reward: [[67.088844]], Actor Loss: [[-12.900847  -44.95893
-3.4410002]], Critic Loss: 3757.785888671875
Episode 86/500, Total Reward: [[84.72262]], Actor Loss: [[-27.501038 -48.645794
-12.483982]], Critic Loss: 7855.4208984375
Episode 87/500, Total Reward: [[215.69992]], Actor Loss: [[ -31.405096
-168.92143    -7.71689 ]], Critic Loss: 43282.0625
Episode 88/500, Total Reward: [[95.90127]], Actor Loss: [[-28.778248 -58.841694
-17.073515]], Critic Loss: 10960.7216796875
Episode 89/500, Total Reward: [[63.567497]], Actor Loss: [[ -9.059072 -58.127365
-6.902358]], Critic Loss: 5489.14892578125
Episode 90/500, Total Reward: [[66.83086]], Actor Loss: [[-24.681656  -38.206444
-4.7035303]], Critic Loss: 4568.62841796875
Episode 91/500, Total Reward: [[74.299644]], Actor Loss: [[-23.058645  -32.72015
-2.9028573]], Critic Loss: 3443.536376953125
Episode 92/500, Total Reward: [[57.643803]], Actor Loss: [[-27.010908 -26.19789
-13.08842 ]], Critic Loss: 4395.3212890625
Episode 93/500, Total Reward: [[128.9641]], Actor Loss: [[-23.513107 -93.03504
-2.263964]], Critic Loss: 14116.3193359375
Episode 94/500, Total Reward: [[56.25491]], Actor Loss: [[-11.546629  -50.187763
-1.5413147]], Critic Loss: 4003.814697265625
Episode 95/500, Total Reward: [[65.23292]], Actor Loss: [[ -9.471692  -50.98105
-1.8015212]], Critic Loss: 3875.593505859375
Episode 96/500, Total Reward: [[72.4478]], Actor Loss: [[ -1.0177807  -59.850334
-0.09336314]], Critic Loss: 3716.302490234375
Episode 97/500, Total Reward: [[76.477776]], Actor Loss: [[-1.4923564e-01
-5.6121811e+01 -5.4530893e-03]], Critic Loss: 3167.044189453125
Episode 98/500, Total Reward: [[83.07009]], Actor Loss: [[ -0.61477894 -84.03767
-0.09734213]], Critic Loss: 7182.5263671875
Episode 99/500, Total Reward: [[66.12926]], Actor Loss: [[ -0.6192976  -73.05931
-0.15991828]], Critic Loss: 5452.12890625
Episode 100/500, Total Reward: [[100.65836]], Actor Loss: [[  -0.4651412
-106.72798     -0.1089745]], Critic Loss: 11513.7392578125
Episode 101/500, Total Reward: [[63.788998]], Actor Loss: [[ -5.8006406
-59.032703   -2.0916765]], Critic Loss: 4478.9580078125
```

```
Episode 102/500, Total Reward: [[103.070564]], Actor Loss: [[ -4.694963
-84.40165    -0.4002927]], Critic Loss: 8009.6953125
Episode 103/500, Total Reward: [[194.69942]], Actor Loss: [[ -20.908106
-169.18115    -1.969904]], Critic Loss: 36886.7265625
Episode 104/500, Total Reward: [[100.17787]], Actor Loss: [[  -1.6862946
-107.899445     -0.21638341]], Critic Loss: 12056.5068359375
Episode 105/500, Total Reward: [[66.81916]], Actor Loss: [[ -1.5298085
-49.57199     -0.11436634]], Critic Loss: 2623.095458984375
Episode 106/500, Total Reward: [[109.11068]], Actor Loss: [[  -2.7305408
-110.68245     -0.5256794]], Critic Loss: 12982.0185546875
Episode 107/500, Total Reward: [[69.88148]], Actor Loss: [[-28.060204 -35.29482
-5.011461]], Critic Loss: 4673.97509765625
Episode 108/500, Total Reward: [[129.47084]], Actor Loss: [[-56.082348 -48.03723
-25.101368]], Critic Loss: 16698.052734375
Episode 109/500, Total Reward: [[59.83401]], Actor Loss: [[-34.866985
-29.142975   -2.9960315]], Critic Loss: 4489.802734375
Episode 110/500, Total Reward: [[93.596504]], Actor Loss: [[-27.640097
-36.882435 -35.64334 ]], Critic Loss: 10033.201171875
Episode 111/500, Total Reward: [[65.47221]], Actor Loss: [[-16.54654  -16.888199
-29.107954]], Critic Loss: 3911.588134765625
Episode 112/500, Total Reward: [[81.676636]], Actor Loss: [[-37.8576
-20.019154  -6.695723]], Critic Loss: 4169.60498046875
Episode 113/500, Total Reward: [[92.99207]], Actor Loss: [[-14.790183  -56.26372
-2.4029825]], Critic Loss: 5395.9140625
Episode 114/500, Total Reward: [[80.94107]], Actor Loss: [[-35.319252 -22.456985
-22.036263]], Critic Loss: 6370.03515625
Episode 115/500, Total Reward: [[85.08374]], Actor Loss: [[-44.42249  -17.235174
-26.054434]], Critic Loss: 7693.41064453125
Episode 116/500, Total Reward: [[70.65712]], Actor Loss: [[-27.945887 -16.880175
-15.027731]], Critic Loss: 3582.4765625
Episode 117/500, Total Reward: [[42.08885]], Actor Loss: [[-13.331761
-14.307831  -13.6727705]], Critic Loss: 1706.7113037109375
Episode 118/500, Total Reward: [[88.061775]], Actor Loss: [[-21.799332
-54.077618 -17.331873]], Critic Loss: 8687.884765625
Episode 119/500, Total Reward: [[65.6701]], Actor Loss: [[-23.894464 -30.157766
-23.406084]], Critic Loss: 5999.7900390625
Episode 120/500, Total Reward: [[65.06572]], Actor Loss: [[-27.584127 -24.33084
-21.527208]], Critic Loss: 5393.75244140625
Episode 121/500, Total Reward: [[44.749485]], Actor Loss: [[ -7.603019
-24.683025  -5.167729]], Critic Loss: 1402.78515625
Episode 122/500, Total Reward: [[95.66314]], Actor Loss: [[-26.852526 -23.288843
-62.516365]], Critic Loss: 12691.763671875
Episode 123/500, Total Reward: [[140.1498]], Actor Loss: [[-62.052708 -53.465385
-20.434162]], Critic Loss: 18483.015625
Episode 124/500, Total Reward: [[118.251976]], Actor Loss: [[-51.5015
-27.543268 -44.201084]], Critic Loss: 15189.5390625
Episode 125/500, Total Reward: [[75.26243]], Actor Loss: [[-27.810253 -36.88957
-15.958135]], Critic Loss: 6505.70654296875
```

```
Episode 126/500, Total Reward: [[116.66184]], Actor Loss: [[-19.698944 -83.11823
-9.611748]], Critic Loss: 12640.2626953125
Episode 127/500, Total Reward: [[108.984825]], Actor Loss: [[-25.106504
-32.447468 -49.564934]], Critic Loss: 11474.4609375
Episode 128/500, Total Reward: [[203.56499]], Actor Loss: [[-51.499928
-63.811043 -92.417496]], Critic Loss: 43151.1171875
Episode 129/500, Total Reward: [[63.16966]], Actor Loss: [[ -8.779012 -15.438829
-32.107056]], Critic Loss: 3172.494140625
Episode 130/500, Total Reward: [[61.357357]], Actor Loss: [[-20.215836
-19.423498 -41.16963 ]], Critic Loss: 6530.087890625
Episode 131/500, Total Reward: [[115.694496]], Actor Loss: [[-23.711977
-48.943607 -40.22871 ]], Critic Loss: 12742.86328125
Episode 132/500, Total Reward: [[194.12834]], Actor Loss: [[-27.998077 -81.55216
-81.105644]], Critic Loss: 36349.66015625
Episode 133/500, Total Reward: [[48.904873]], Actor Loss: [[-29.096636
-9.476079 -14.920439]], Critic Loss: 2861.517333984375
Episode 134/500, Total Reward: [[67.89178]], Actor Loss: [[-38.666782 -24.777323
-14.1537  ]], Critic Loss: 6021.419921875
Episode 135/500, Total Reward: [[56.663204]], Actor Loss: [[-18.653717
-20.376366 -19.194084]], Critic Loss: 3390.0537109375
Episode 136/500, Total Reward: [[42.687366]], Actor Loss: [[-16.183369
-12.912292 -11.231153]], Critic Loss: 1626.2518310546875
Episode 137/500, Total Reward: [[100.28884]], Actor Loss: [[-35.62672
-34.632507 -30.707968]], Critic Loss: 10194.3740234375
Episode 138/500, Total Reward: [[128.80737]], Actor Loss: [[-51.42795
-54.885395 -24.243898]], Critic Loss: 17045.1953125
Episode 139/500, Total Reward: [[92.21496]], Actor Loss: [[-39.53648  -26.234772
-21.513762]], Critic Loss: 7618.67431640625
Episode 140/500, Total Reward: [[77.60207]], Actor Loss: [[-41.916157 -16.564705
-16.197195]], Critic Loss: 5576.81201171875
Episode 141/500, Total Reward: [[136.35373]], Actor Loss: [[-67.40584
-50.380558 -27.329723]], Critic Loss: 21058.68359375
Episode 142/500, Total Reward: [[42.918602]], Actor Loss: [[-14.18949
-8.329694 -16.03736 ]], Critic Loss: 1486.607177734375
Episode 143/500, Total Reward: [[93.65446]], Actor Loss: [[-18.516851 -19.097698
-55.200817]], Critic Loss: 8614.6923828125
Episode 144/500, Total Reward: [[81.083374]], Actor Loss: [[-32.39598
-15.624732 -32.68421 ]], Critic Loss: 6513.28515625
Episode 145/500, Total Reward: [[116.75259]], Actor Loss: [[-21.442696 -53.33624
-26.895193]], Critic Loss: 10337.6279296875
Episode 146/500, Total Reward: [[54.14188]], Actor Loss: [[-18.698446 -10.63643
-33.093204]], Critic Loss: 3897.264892578125
Episode 147/500, Total Reward: [[164.30351]], Actor Loss: [[-43.769825 -94.1331
-9.310276]], Critic Loss: 21671.724609375
Episode 148/500, Total Reward: [[208.24431]], Actor Loss: [[ -76.13636
-113.017914  -11.558992]], Critic Loss: 40285.8125
Episode 149/500, Total Reward: [[49.471294]], Actor Loss: [[-23.798729
-12.488938 -22.79148 ]], Critic Loss: 3490.345703125
```

```
Episode 150/500, Total Reward: [[57.342144]], Actor Loss: [[-27.212502
-13.690133 -22.804651]], Critic Loss: 4058.618408203125
Episode 151/500, Total Reward: [[139.61156]], Actor Loss: [[-59.190636
-21.580462 -56.666397]], Critic Loss: 18889.06640625
Episode 152/500, Total Reward: [[94.859825]], Actor Loss: [[-35.574604
-32.215633 -11.657718]], Critic Loss: 6311.978515625
Episode 153/500, Total Reward: [[67.57706]], Actor Loss: [[-33.516895 -13.437421
-37.828773]], Critic Loss: 7188.17236328125
Episode 154/500, Total Reward: [[117.746056]], Actor Loss: [[-75.76337
-25.733221 -15.511653]], Critic Loss: 13690.927734375
Episode 155/500, Total Reward: [[149.2994]], Actor Loss: [[-66.930466 -55.500698
-18.120071]], Critic Loss: 19754.650390625
Episode 156/500, Total Reward: [[69.71345]], Actor Loss: [[-38.796844 -17.348425
-16.519083]], Critic Loss: 5280.10791015625
Episode 157/500, Total Reward: [[80.668144]], Actor Loss: [[-48.806854
-10.568882  -15.1555605]], Critic Loss: 5554.9140625
Episode 158/500, Total Reward: [[80.668144]], Actor Loss: [[-64.20448
-13.903162 -19.93685 ]], Critic Loss: 9612.72265625
Episode 159/500, Total Reward: [[69.57722]], Actor Loss: [[-41.53006  -18.95045
-21.659708]], Critic Loss: 6747.015625
Episode 160/500, Total Reward: [[81.978874]], Actor Loss: [[-53.567963
-11.886369 -12.174967]], Critic Loss: 6026.3076171875
Episode 161/500, Total Reward: [[54.65459]], Actor Loss: [[-35.91554  -17.72255
-10.300676]], Critic Loss: 4088.165771484375
Episode 162/500, Total Reward: [[62.322765]], Actor Loss: [[-43.16778
-19.989258 -13.772511]], Critic Loss: 5918.15576171875
Episode 163/500, Total Reward: [[72.34578]], Actor Loss: [[-43.131836
-12.274698  -12.1322365]], Critic Loss: 4561.48583984375
Episode 164/500, Total Reward: [[109.58379]], Actor Loss: [[-43.622948
-45.513344  -7.304782]], Critic Loss: 9300.8798828125
Episode 165/500, Total Reward: [[116.927185]], Actor Loss: [[-61.772636
-21.684214 -32.18907 ]], Critic Loss: 13373.98046875
Episode 166/500, Total Reward: [[93.35113]], Actor Loss: [[-62.38787  -15.309304
-20.964668]], Critic Loss: 9734.1591796875
Episode 167/500, Total Reward: [[120.90236]], Actor Loss: [[-100.16318
-15.57341   -22.821417]], Critic Loss: 19198.318359375
Episode 168/500, Total Reward: [[95.292656]], Actor Loss: [[-65.62508
-15.038177   -3.1103845]], Critic Loss: 7018.02197265625
Episode 169/500, Total Reward: [[129.1006]], Actor Loss: [[-62.640705  -65.28603
-4.5208106]], Critic Loss: 17542.35546875
Episode 170/500, Total Reward: [[61.399853]], Actor Loss: [[-35.700264
-10.678137 -18.623568]], Critic Loss: 4225.255859375
Episode 171/500, Total Reward: [[122.52436]], Actor Loss: [[-67.243546
-38.229614   -3.8941357]], Critic Loss: 11961.203125
Episode 172/500, Total Reward: [[124.53235]], Actor Loss: [[-38.0416
-42.111423 -34.01089 ]], Critic Loss: 13033.400390625
Episode 173/500, Total Reward: [[96.1508]], Actor Loss: [[-35.11949  -21.06889
-49.769356]], Critic Loss: 11227.041015625
```

```
Episode 174/500, Total Reward: [[69.51459]], Actor Loss: [[-29.908041 -15.498149
-40.863033]], Critic Loss: 7442.37939453125
Episode 175/500, Total Reward: [[131.57793]], Actor Loss: [[  -8.107531
-104.32634    -6.650293]], Critic Loss: 14181.037109375
Episode 176/500, Total Reward: [[59.84477]], Actor Loss: [[-21.569845 -15.563463
-16.824955]], Critic Loss: 2911.49462890625
Episode 177/500, Total Reward: [[174.56387]], Actor Loss: [[ -22.99027
-117.88887   -17.289444]], Critic Loss: 25017.298828125
Episode 178/500, Total Reward: [[73.12965]], Actor Loss: [[-23.30479  -14.980669
-27.25365 ]], Critic Loss: 4295.37451171875
Episode 179/500, Total Reward: [[126.96329]], Actor Loss: [[-16.350769 -79.27369
-15.533334]], Critic Loss: 12356.056640625
Episode 180/500, Total Reward: [[59.047432]], Actor Loss: [[ -9.657994 -38.97747
-17.686584]], Critic Loss: 4398.61474609375
Episode 181/500, Total Reward: [[56.68951]], Actor Loss: [[-13.712955 -33.267185
-19.576262]], Critic Loss: 4429.7548828125
Episode 182/500, Total Reward: [[120.174324]], Actor Loss: [[ -8.484055
-80.04965  -13.44596 ]], Critic Loss: 10399.8525390625
Episode 183/500, Total Reward: [[62.54498]], Actor Loss: [[-10.700049 -32.32442
-15.12506 ]], Critic Loss: 3381.36767578125
Episode 184/500, Total Reward: [[68.58056]], Actor Loss: [[-16.303768 -16.171137
-22.774097]], Critic Loss: 3052.4521484375
Episode 185/500, Total Reward: [[112.02383]], Actor Loss: [[-60.904564
-27.899448 -47.167885]], Critic Loss: 18488.35546875
Episode 186/500, Total Reward: [[40.25579]], Actor Loss: [[-10.426909
-8.440249   -6.9668164]], Critic Loss: 667.3943481445312
Episode 187/500, Total Reward: [[76.68203]], Actor Loss: [[-21.169666 -57.26127
-10.590947]], Critic Loss: 7924.896484375
Episode 188/500, Total Reward: [[60.29755]], Actor Loss: [[-19.55508  -25.556955
-6.07361 ]], Critic Loss: 2619.970458984375
Episode 189/500, Total Reward: [[78.257416]], Actor Loss: [[-32.380993
-33.552532 -18.501303]], Critic Loss: 7129.240234375
Episode 190/500, Total Reward: [[83.65084]], Actor Loss: [[-51.220184 -27.149525
-17.324137]], Critic Loss: 9157.3125
Episode 191/500, Total Reward: [[128.42787]], Actor Loss: [[-18.59922
-86.57516    -1.1750585]], Critic Loss: 11310.2041015625
Episode 192/500, Total Reward: [[79.4643]], Actor Loss: [[-11.316973  -64.13113
-1.4060237]], Critic Loss: 5906.556640625
Episode 193/500, Total Reward: [[115.53764]], Actor Loss: [[-6.6049933e-01
-1.1398832e+02 -7.1741022e-02]], Critic Loss: 13160.806640625
Episode 194/500, Total Reward: [[158.5136]], Actor Loss: [[  -6.0640755
-154.54332      -0.58127046]], Critic Loss: 25981.783203125
Episode 195/500, Total Reward: [[60.830902]], Actor Loss: [[ -5.550421
-60.089306   -0.5689129]], Critic Loss: 4383.583984375
Episode 196/500, Total Reward: [[119.6434]], Actor Loss: [[  -3.8655303
-118.79954     -0.4903739]], Critic Loss: 15167.2646484375
Episode 197/500, Total Reward: [[61.956707]], Actor Loss: [[-4.8947621e-02
-5.2243374e+01 -6.5992004e-03]], Critic Loss: 2735.177001953125
```

```
Episode 198/500, Total Reward: [[89.07478]], Actor Loss: [[-1.0300527e-01
-8.8938217e+01 -6.4274031e-03]], Critic Loss: 7929.4833984375
Episode 199/500, Total Reward: [[74.64648]], Actor Loss: [[-3.4322160e-01
-9.2808914e+01 -4.3482378e-02]], Critic Loss: 8685.421875
Episode 200/500, Total Reward: [[100.17356]], Actor Loss: [[  -2.2749634
-108.89194      -0.30430403]], Critic Loss: 12425.828125
Episode 201/500, Total Reward: [[207.6046]], Actor Loss: [[ -48.197083 -142.89
-5.342845]], Critic Loss: 38584.71875
Episode 202/500, Total Reward: [[93.959526]], Actor Loss: [[-33.55947
-54.415234   -3.2579124]], Critic Loss: 8323.3896484375
Episode 203/500, Total Reward: [[87.60524]], Actor Loss: [[-27.28028
-55.544106   -3.4238844]], Critic Loss: 7438.763671875
Episode 204/500, Total Reward: [[115.230736]], Actor Loss: [[-25.481094
-65.97422    -1.8756238]], Critic Loss: 8710.6640625
Episode 205/500, Total Reward: [[170.70436]], Actor Loss: [[ -25.947279
-132.6126     -1.599424]], Critic Loss: 25651.001953125
Episode 206/500, Total Reward: [[50.15539]], Actor Loss: [[-19.31159   -30.24704
-1.8951505]], Critic Loss: 2647.49169921875
Episode 207/500, Total Reward: [[61.94721]], Actor Loss: [[ -3.4153588
-63.639576    -0.28021848]], Critic Loss: 4534.0224609375
Episode 208/500, Total Reward: [[234.83177]], Actor Loss: [[-6.3072267e+00
-1.9820799e+02 -1.4640664e-01]], Critic Loss: 41886.37890625
Episode 209/500, Total Reward: [[85.68093]], Actor Loss: [[-22.92907   -69.63403
-3.4709473]], Critic Loss: 9222.5400390625
Episode 210/500, Total Reward: [[120.291084]], Actor Loss: [[  -3.1903138
-117.16209      -0.22496726]], Critic Loss: 14538.900390625
Episode 211/500, Total Reward: [[51.28611]], Actor Loss: [[-35.04808  -24.666918
-4.103789]], Critic Loss: 4072.837646484375
Episode 212/500, Total Reward: [[100.596825]], Actor Loss: [[-40.535126
-65.49154    -4.2682643]], Critic Loss: 12164.9716796875
Episode 213/500, Total Reward: [[109.86184]], Actor Loss: [[-18.849007  -77.565
-1.0733235]], Critic Loss: 9503.7802734375
Episode 214/500, Total Reward: [[58.087967]], Actor Loss: [[-34.707138
-32.638218   -2.5863426]], Critic Loss: 4890.44287109375
Episode 215/500, Total Reward: [[293.1053]], Actor Loss: [[ -51.978848
-238.85895     -0.5560064]], Critic Loss: 84910.34375
Episode 216/500, Total Reward: [[175.93114]], Actor Loss: [[-86.6835
-85.21306    -1.4690914]], Critic Loss: 30055.646484375
Episode 217/500, Total Reward: [[83.58913]], Actor Loss: [[-43.073387  -52.91578
-0.5781621]], Critic Loss: 9325.2490234375
Episode 218/500, Total Reward: [[92.84482]], Actor Loss: [[-28.972567
-63.239464    -0.26732862]], Critic Loss: 8552.4326171875
Episode 219/500, Total Reward: [[67.9211]], Actor Loss: [[-16.527018   -55.57309
-0.21223618]], Critic Loss: 5229.07470703125
Episode 220/500, Total Reward: [[258.5955]], Actor Loss: [[-1.6604288e+01
-2.2935736e+02 -8.3680078e-02]], Critic Loss: 60538.3125
Episode 221/500, Total Reward: [[44.686268]], Actor Loss: [[-7.5874579e-01
-3.9813519e+01 -1.2494177e-02]], Critic Loss: 1647.122314453125
```

```
Episode 222/500, Total Reward: [[66.17819]], Actor Loss: [[-7.159397e+00
-6.572900e+01 -5.052116e-02]], Critic Loss: 5320.0859375
Episode 223/500, Total Reward: [[195.67317]], Actor Loss: [[-2.5554023e+00
-1.7423181e+02 -1.0033459e-02]], Critic Loss: 31257.263671875
Episode 224/500, Total Reward: [[84.694984]], Actor Loss: [[-48.778683
-24.626678    -0.25177032]], Critic Loss: 5425.37353515625
Episode 225/500, Total Reward: [[141.54457]], Actor Loss: [[-90.57913
-49.917408   -0.6919668]], Critic Loss: 19934.1953125
Episode 226/500, Total Reward: [[97.59733]], Actor Loss: [[-51.992455
-27.719292    -0.54311955]], Critic Loss: 6440.8427734375
Episode 227/500, Total Reward: [[54.18243]], Actor Loss: [[-20.83978   -18.63526
-0.6448346]], Critic Loss: 1609.604248046875
Episode 228/500, Total Reward: [[58.654823]], Actor Loss: [[-35.93391  -55.42211
-8.215405]], Critic Loss: 9914.46875
Episode 229/500, Total Reward: [[78.32087]], Actor Loss: [[-52.990276
-38.371445   -3.8991823]], Critic Loss: 9074.6396484375
Episode 230/500, Total Reward: [[87.912415]], Actor Loss: [[-45.67915
-41.879864 -13.800271]], Critic Loss: 10273.7041015625
Episode 231/500, Total Reward: [[149.36382]], Actor Loss: [[-56.717094 -70.12829
-30.161026]], Critic Loss: 24651.01171875
Episode 232/500, Total Reward: [[69.04002]], Actor Loss: [[-30.590816 -32.458027
-17.0439  ]], Critic Loss: 6414.84765625
Episode 233/500, Total Reward: [[145.50749]], Actor Loss: [[-50.633022 -65.78456
-28.69139 ]], Critic Loss: 21056.615234375
Episode 234/500, Total Reward: [[114.26461]], Actor Loss: [[-36.274467
-48.624046 -12.019971]], Critic Loss: 9393.19140625
Episode 235/500, Total Reward: [[158.32002]], Actor Loss: [[-73.93128  -46.85253
-37.471912]], Critic Loss: 25044.87890625
Episode 236/500, Total Reward: [[243.04298]], Actor Loss: [[-140.39046
-47.199642  -41.5122  ]], Critic Loss: 52487.85546875
Episode 237/500, Total Reward: [[98.6802]], Actor Loss: [[-13.2181635 -41.910202
-58.79232  ]], Critic Loss: 12977.9228515625
Episode 238/500, Total Reward: [[163.87733]], Actor Loss: [[-30.619854
-51.052593 -74.87981 ]], Critic Loss: 24508.611328125
Episode 239/500, Total Reward: [[63.321663]], Actor Loss: [[ -0.7982009
-25.812515  -16.24587  ]], Critic Loss: 1836.6873779296875
Episode 240/500, Total Reward: [[75.83053]], Actor Loss: [[ -0.8037284
-27.330101  -38.30545  ]], Critic Loss: 4414.177734375
Episode 241/500, Total Reward: [[89.996544]], Actor Loss: [[ -0.7728773
-22.869556  -54.97867  ]], Critic Loss: 6181.2763671875
Episode 242/500, Total Reward: [[223.58583]], Actor Loss: [[ -20.717491
-112.33069   -88.93067 ]], Critic Loss: 49274.609375
Episode 243/500, Total Reward: [[202.64891]], Actor Loss: [[ -19.480476
-150.0815    -45.164795]], Critic Loss: 46107.58984375
Episode 244/500, Total Reward: [[114.00288]], Actor Loss: [[ -1.6616535
-83.90833    -7.9094467]], Critic Loss: 8738.404296875
Episode 245/500, Total Reward: [[44.698933]], Actor Loss: [[ -3.1101944
-17.71588   -18.412687 ]], Critic Loss: 1539.680419921875
```

```
Episode 246/500, Total Reward: [[129.39157]], Actor Loss: [[-11.431843
-82.391464 -40.207043]], Critic Loss: 17964.134765625
Episode 247/500, Total Reward: [[52.961143]], Actor Loss: [[ -5.505124
-32.919327 -24.159565]], Critic Loss: 3916.759033203125
Episode 248/500, Total Reward: [[58.933807]], Actor Loss: [[ -7.254486
-40.639606 -20.092045]], Critic Loss: 4622.11474609375
Episode 249/500, Total Reward: [[70.561386]], Actor Loss: [[ -6.7248673
-28.869543  -20.398022 ]], Critic Loss: 3135.15234375
Episode 250/500, Total Reward: [[156.49648]], Actor Loss: [[  -8.152111
-111.0182    -32.23697 ]], Critic Loss: 22924.166015625
Episode 251/500, Total Reward: [[78.15018]], Actor Loss: [[-11.1069565
-40.762794  -28.014683 ]], Critic Loss: 6381.5234375
Episode 252/500, Total Reward: [[96.476006]], Actor Loss: [[ -0.6508057
-93.78828   -10.887102 ]], Critic Loss: 11093.6044921875
Episode 253/500, Total Reward: [[144.16231]], Actor Loss: [[  -0.52531
-123.84223     -4.2602267]], Critic Loss: 16545.10546875
Episode 254/500, Total Reward: [[85.00811]], Actor Loss: [[-9.9411106e-04
-6.5717606e+01 -1.4932592e-02]], Critic Loss: 4320.89794921875
Episode 255/500, Total Reward: [[69.53123]], Actor Loss: [[-3.3715546e-02
-9.1838921e+01 -6.0345024e-01]], Critic Loss: 8551.8271484375
Episode 256/500, Total Reward: [[63.050022]], Actor Loss: [[-3.2115029e-05
-4.4103214e+01 -1.5800165e-03]], Critic Loss: 1945.23583984375
Episode 257/500, Total Reward: [[186.60658]], Actor Loss: [[-5.9000868e-04
-2.1604738e+02 -6.0121138e-02]], Critic Loss: 46702.70703125
Episode 258/500, Total Reward: [[118.35133]], Actor Loss: [[-3.9613882e-05
-1.0080023e+02 -1.9463365e-03]], Critic Loss: 10161.0869140625
Episode 259/500, Total Reward: [[138.27101]], Actor Loss: [[-2.7091410e-03
-1.6951570e+02 -1.1313009e-01]], Critic Loss: 28774.85546875
Episode 260/500, Total Reward: [[98.93081]], Actor Loss: [[-8.9213578e-03
-9.7288933e+01 -1.4051157e-01]], Critic Loss: 9494.2333984375
Episode 261/500, Total Reward: [[166.82808]], Actor Loss: [[-5.287332e-03
-1.490497e+02 -9.897585e-02]], Critic Loss: 22246.904296875
Episode 262/500, Total Reward: [[62.372124]], Actor Loss: [[-4.8946172e-02
-7.0896988e+01 -1.0903988e+00]], Critic Loss: 5189.23291015625
Episode 263/500, Total Reward: [[81.62497]], Actor Loss: [[-5.2177183e-02
-9.4388420e+01 -1.3025353e+00]], Critic Loss: 9166.748046875
Episode 264/500, Total Reward: [[105.55016]], Actor Loss: [[-1.4497908e-02
-8.9043724e+01 -3.4690875e-01]], Critic Loss: 7993.27685546875
Episode 265/500, Total Reward: [[55.55572]], Actor Loss: [[-1.6629146e-02
-5.5727337e+01 -3.6932063e-01]], Critic Loss: 3148.701171875
Episode 266/500, Total Reward: [[55.628567]], Actor Loss: [[-1.2368379e-03
-3.6564857e+01 -4.2618442e-02]], Critic Loss: 1340.19775390625
Episode 267/500, Total Reward: [[97.96585]], Actor Loss: [[-3.61477793e-03
-9.66000366e+01 -1.15568854e-01]], Critic Loss: 9354.6083984375
Episode 268/500, Total Reward: [[99.93221]], Actor Loss: [[-1.8669323e-04
-9.9294174e+01 -4.8485887e-03]], Critic Loss: 9860.3349609375
Episode 269/500, Total Reward: [[90.394104]], Actor Loss: [[-1.7797494e-04
-8.2712708e+01 -6.2095993e-03]], Critic Loss: 6842.44970703125
```

```
Episode 270/500, Total Reward: [[78.76067]], Actor Loss: [[-7.4799086e-06
-6.0046478e+01 -3.5418788e-04]], Critic Loss: 3605.62255859375
Episode 271/500, Total Reward: [[100.55077]], Actor Loss: [[-5.0358347e-05
-9.8172112e+01 -1.6396943e-03]], Critic Loss: 9638.0947265625
Episode 272/500, Total Reward: [[85.92673]], Actor Loss: [[-3.14263999e-03
-8.23887482e+01 -1.20357975e-01]], Critic Loss: 6808.27197265625
Episode 273/500, Total Reward: [[44.58817]], Actor Loss: [[-4.1749193e-03
-7.9975159e+01 -1.4722981e-01]], Critic Loss: 6420.26611328125
Episode 274/500, Total Reward: [[48.97733]], Actor Loss: [[-4.1575604e-03
-5.7940254e+01 -1.6414478e-01]], Critic Loss: 3376.6044921875
Episode 275/500, Total Reward: [[70.71709]], Actor Loss: [[-7.5229928e-03
-1.0320164e+02 -1.8769543e-01]], Critic Loss: 10690.9091796875
Episode 276/500, Total Reward: [[49.855373]], Actor Loss: [[-2.2843398e-02
-5.0037411e+01 -5.1187211e-01]], Critic Loss: 2557.5400390625
Episode 277/500, Total Reward: [[61.18762]], Actor Loss: [[-3.4661076e-03
-5.6239506e+01 -1.1067373e-01]], Critic Loss: 3175.7333984375
Episode 278/500, Total Reward: [[47.400703]], Actor Loss: [[-6.138918e-03
-5.127674e+01 -2.412594e-01]], Critic Loss: 2654.737060546875
Episode 279/500, Total Reward: [[135.87584]], Actor Loss: [[-4.9990811e-03
-1.4016194e+02 -2.0940296e-01]], Critic Loss: 19705.517578125
Episode 280/500, Total Reward: [[57.40817]], Actor Loss: [[-2.9158157e-03
-5.9742676e+01 -1.3939004e-01]], Critic Loss: 3586.210693359375
Episode 281/500, Total Reward: [[96.476006]], Actor Loss: [[-2.1432537e-04
-7.5770287e+01 -6.2077311e-03]], Critic Loss: 5742.10986328125
Episode 282/500, Total Reward: [[94.75177]], Actor Loss: [[-1.0059610e-03
-9.6732048e+01 -1.4926775e-02]], Critic Loss: 9360.1708984375
Episode 283/500, Total Reward: [[61.265675]], Actor Loss: [[-7.1242318e-04
-6.7565254e+01 -2.4281563e-02]], Critic Loss: 4568.44140625
Episode 284/500, Total Reward: [[64.17723]], Actor Loss: [[-3.2995983e-03
-6.0154007e+01 -5.2771881e-02]], Critic Loss: 3625.253173828125
Episode 285/500, Total Reward: [[212.97153]], Actor Loss: [[-1.3243093e-01
-1.9575171e+02 -4.6506248e+00]], Critic Loss: 40214.19140625
Episode 286/500, Total Reward: [[102.27853]], Actor Loss: [[-1.5640588e-02
-9.5395515e+01 -2.8612310e-01]], Critic Loss: 9157.9677734375
Episode 287/500, Total Reward: [[55.536915]], Actor Loss: [[ -0.4581938
-57.50761    -7.828765 ]], Critic Loss: 4328.92578125
Episode 288/500, Total Reward: [[83.72046]], Actor Loss: [[-1.1625452e-02
-6.1396767e+01 -8.5523225e-02]], Critic Loss: 3781.501220703125
Episode 289/500, Total Reward: [[164.9092]], Actor Loss: [[  -7.7620754
-109.32599    -70.266495 ]], Critic Loss: 35101.73046875
Episode 290/500, Total Reward: [[77.07397]], Actor Loss: [[ -1.2590038 -56.76342
-11.716301 ]], Critic Loss: 4863.48974609375
Episode 291/500, Total Reward: [[93.84869]], Actor Loss: [[ -2.276919 -56.361374
-37.062653]], Critic Loss: 9158.671875
Episode 292/500, Total Reward: [[75.97868]], Actor Loss: [[-3.6885895e-02
-6.0435177e+01 -1.7750971e+00]], Critic Loss: 3874.709228515625
Episode 293/500, Total Reward: [[72.353294]], Actor Loss: [[ -0.07900412
-70.93256     -7.498142  ]], Critic Loss: 6163.77392578125
```

```
Episode 294/500, Total Reward: [[110.18317]], Actor Loss: [[-2.4370139e-03
-1.0197012e+02 -5.6976721e-02]], Critic Loss: 10410.0244140625
Episode 295/500, Total Reward: [[57.825848]], Actor Loss: [[ -1.0764415
-40.401127  -36.10545  ]], Critic Loss: 6019.12451171875
Episode 296/500, Total Reward: [[72.352844]], Actor Loss: [[ -3.3994684
-49.317554  -26.317238 ]], Critic Loss: 6246.4150390625
Episode 297/500, Total Reward: [[62.226994]], Actor Loss: [[ -0.50041085
-64.024826    -8.155456  ]], Critic Loss: 5282.482421875
Episode 298/500, Total Reward: [[56.124176]], Actor Loss: [[ -0.5516846
-34.452023  -10.063807 ]], Critic Loss: 2031.0806884765625
Episode 299/500, Total Reward: [[212.33766]], Actor Loss: [[  -2.5528946
-156.06805    -53.913795 ]], Critic Loss: 45171.01171875
Episode 300/500, Total Reward: [[56.828323]], Actor Loss: [[ -1.4261835
-28.021465  -29.658073 ]], Critic Loss: 3493.486083984375
Episode 301/500, Total Reward: [[69.59376]], Actor Loss: [[ -1.534631 -27.698256
-20.187435]], Critic Loss: 2442.367919921875
Episode 302/500, Total Reward: [[93.425095]], Actor Loss: [[ -2.7760565
-56.155323  -22.881477 ]], Critic Loss: 6693.34375
Episode 303/500, Total Reward: [[66.30417]], Actor Loss: [[ -6.2324624
-24.588528  -36.45219  ]], Critic Loss: 4525.6806640625
Episode 304/500, Total Reward: [[72.22895]], Actor Loss: [[ -3.1661575
-40.022453  -16.64553  ]], Critic Loss: 3580.124755859375
Episode 305/500, Total Reward: [[94.36776]], Actor Loss: [[ -0.6752465 -77.27585
-6.211919 ]], Critic Loss: 7083.41357421875
Episode 306/500, Total Reward: [[140.54037]], Actor Loss: [[ -4.0790715
-92.77322   -39.62769  ]], Critic Loss: 18626.78515625
Episode 307/500, Total Reward: [[167.4875]], Actor Loss: [[  -0.98362315
-150.48015      -7.242788  ]], Critic Loss: 25187.77734375
Episode 308/500, Total Reward: [[47.714844]], Actor Loss: [[-2.3500780e-02
-4.0522266e+01 -5.1926553e-01]], Critic Loss: 1686.3369140625
Episode 309/500, Total Reward: [[54.070415]], Actor Loss: [[ -0.3730355
-69.45812    -6.4363923]], Critic Loss: 5816.73876953125
Episode 310/500, Total Reward: [[108.00172]], Actor Loss: [[ -1.8252993
-83.70386   -14.138902 ]], Critic Loss: 9933.720703125
Episode 311/500, Total Reward: [[60.784893]], Actor Loss: [[ -6.958135
-15.398967 -39.139366]], Critic Loss: 3781.8154296875
Episode 312/500, Total Reward: [[100.914604]], Actor Loss: [[ -8.866538
-54.482338 -36.91561 ]], Critic Loss: 10052.96875
Episode 313/500, Total Reward: [[101.83231]], Actor Loss: [[ -1.2272074
-32.323425  -66.450294 ]], Critic Loss: 10000.1845703125
Episode 314/500, Total Reward: [[167.68607]], Actor Loss: [[  -6.3262954
-157.29662    -10.6111555]], Critic Loss: 30357.51171875
Episode 315/500, Total Reward: [[58.96466]], Actor Loss: [[ -6.8353753
-46.067802  -14.136932 ]], Critic Loss: 4494.376953125
Episode 316/500, Total Reward: [[68.17422]], Actor Loss: [[ -1.9568971
-72.891426   -1.9103465]], Critic Loss: 5891.89404296875
Episode 317/500, Total Reward: [[72.137474]], Actor Loss: [[ -1.7397869
-73.53947    -5.0961266]], Critic Loss: 6460.2021484375
```

```
Episode 318/500, Total Reward: [[80.69917]], Actor Loss: [[ -1.2463673 -81.05763
-2.313008 ]], Critic Loss: 7160.03857421875
Episode 319/500, Total Reward: [[103.75812]], Actor Loss: [[  -0.5016008
-117.239975    -0.8023455]], Critic Loss: 14052.6611328125
Episode 320/500, Total Reward: [[92.23706]], Actor Loss: [[ -0.10019235
-70.10792     -0.1645793 ]], Critic Loss: 4952.3154296875
Episode 321/500, Total Reward: [[186.56497]], Actor Loss: [[  -1.3949523
-184.87837     -2.02148  ]], Critic Loss: 35454.9296875
Episode 322/500, Total Reward: [[84.48372]], Actor Loss: [[-13.100587 -51.37907
-18.696209]], Critic Loss: 6918.224609375
Episode 323/500, Total Reward: [[110.12527]], Actor Loss: [[-47.908836
-28.638735 -46.513683]], Critic Loss: 15144.0712890625
Episode 324/500, Total Reward: [[53.317123]], Actor Loss: [[-15.307275
-9.785907 -20.59434 ]], Critic Loss: 2087.349853515625
Episode 325/500, Total Reward: [[73.82882]], Actor Loss: [[-32.06041  -29.252798
-21.700823]], Critic Loss: 6891.3291015625
Episode 326/500, Total Reward: [[98.734764]], Actor Loss: [[ -3.5127947
-83.58293    -4.162653 ]], Critic Loss: 8328.0927734375
Episode 327/500, Total Reward: [[110.207985]], Actor Loss: [[ -9.3498335
-99.699844   -8.969885 ]], Critic Loss: 13928.6171875
Episode 328/500, Total Reward: [[57.617954]], Actor Loss: [[ -7.305189
-49.485043  -10.6644335]], Critic Loss: 4550.1318359375
Episode 329/500, Total Reward: [[63.329346]], Actor Loss: [[ -2.9293258
-51.801846   -4.836438 ]], Critic Loss: 3548.30029296875
Episode 330/500, Total Reward: [[65.03092]], Actor Loss: [[ -2.1689517
-55.235874   -2.7774065]], Critic Loss: 3621.900634765625
Episode 331/500, Total Reward: [[109.356575]], Actor Loss: [[-14.544983
-73.713425 -20.779308]], Critic Loss: 11889.22265625
Episode 332/500, Total Reward: [[120.52545]], Actor Loss: [[ -5.3740478 -91.9105
-3.2567933]], Critic Loss: 10108.560546875
Episode 333/500, Total Reward: [[60.507107]], Actor Loss: [[-19.437603
-20.562334 -17.451277]], Critic Loss: 3300.642333984375
Episode 334/500, Total Reward: [[70.156364]], Actor Loss: [[-14.5481415
-49.945038  -10.381383 ]], Critic Loss: 5606.20068359375
Episode 335/500, Total Reward: [[135.90341]], Actor Loss: [[-47.205143
-54.155655 -35.17908 ]], Critic Loss: 18643.13671875
Episode 336/500, Total Reward: [[57.49337]], Actor Loss: [[-20.689486 -20.505587
-16.728529]], Critic Loss: 3355.143310546875
Episode 337/500, Total Reward: [[148.95625]], Actor Loss: [[-43.42162
-55.427254 -39.774548]], Critic Loss: 19216.455078125
Episode 338/500, Total Reward: [[186.50548]], Actor Loss: [[-52.97565 -68.66977
-60.26526]], Critic Loss: 33091.4921875
Episode 339/500, Total Reward: [[59.63434]], Actor Loss: [[-10.55317  -18.20198
-4.013822]], Critic Loss: 1073.8056640625
Episode 340/500, Total Reward: [[95.208244]], Actor Loss: [[-14.9734125
-69.83484    -8.10084  ]], Critic Loss: 8632.09765625
Episode 341/500, Total Reward: [[85.37589]], Actor Loss: [[-19.303844  -55.74936
-2.6230137]], Critic Loss: 6033.5947265625
```

```
Episode 342/500, Total Reward: [[49.661533]], Actor Loss: [[-19.225163
-1.6991074 -24.556269 ]], Critic Loss: 2068.479248046875
Episode 343/500, Total Reward: [[85.67]], Actor Loss: [[-50.774128   -2.5249314
-33.55744  ]], Critic Loss: 7544.052734375
Episode 344/500, Total Reward: [[141.54475]], Actor Loss: [[-77.657585
-2.8816059 -56.4853   ]], Critic Loss: 18775.7109375
Episode 345/500, Total Reward: [[92.23706]], Actor Loss: [[-47.03131   -8.025004
-19.082382]], Critic Loss: 5496.5458984375
Episode 346/500, Total Reward: [[83.980995]], Actor Loss: [[-24.737047
-10.873828 -49.779427]], Critic Loss: 7291.50390625
Episode 347/500, Total Reward: [[83.62786]], Actor Loss: [[-17.718159 -16.680838
-48.351223]], Critic Loss: 6847.59912109375
Episode 348/500, Total Reward: [[87.93458]], Actor Loss: [[ -8.041025
-7.0974255 -87.69134  ]], Critic Loss: 10573.966796875
Episode 349/500, Total Reward: [[159.61806]], Actor Loss: [[  -1.0313604
-5.4934244 -141.178    ]], Critic Loss: 21816.109375
Episode 350/500, Total Reward: [[74.78763]], Actor Loss: [[ -0.16907188
-7.591659   -47.32015   ]], Critic Loss: 3033.9033203125
Episode 351/500, Total Reward: [[55.410347]], Actor Loss: [[ -0.6058164
-1.6551955 -56.953995 ]], Critic Loss: 3506.41748046875
Episode 352/500, Total Reward: [[75.78956]], Actor Loss: [[-3.4216259e-02
-3.2440734e-01 -6.6062386e+01]], Critic Loss: 4411.75
Episode 353/500, Total Reward: [[66.645485]], Actor Loss: [[ -1.1807561
-1.1105514 -61.045414 ]], Critic Loss: 4011.54052734375
Episode 354/500, Total Reward: [[79.696815]], Actor Loss: [[ -0.9523654
-1.456175  -71.94716  ]], Critic Loss: 5528.76953125
Episode 355/500, Total Reward: [[112.79504]], Actor Loss: [[  -2.410665
-2.2745562 -102.08981  ]], Critic Loss: 11400.9091796875
Episode 356/500, Total Reward: [[46.35742]], Actor Loss: [[ -0.24349633
-0.32537264 -64.88971   ]], Critic Loss: 4284.82470703125
Episode 357/500, Total Reward: [[76.09339]], Actor Loss: [[ -1.0459851
-1.0034916 -67.82361  ]], Critic Loss: 4882.248046875
Episode 358/500, Total Reward: [[228.69836]], Actor Loss: [[  -2.7254925
-9.003528  -190.27014  ]], Critic Loss: 40803.66796875
Episode 359/500, Total Reward: [[58.329605]], Actor Loss: [[ -0.0744893
-0.24411196 -68.97318   ]], Critic Loss: 4801.3515625
Episode 360/500, Total Reward: [[100.41999]], Actor Loss: [[-4.5440465e-02
-2.0584515e-01 -8.6214279e+01]], Critic Loss: 7476.29443359375
Episode 361/500, Total Reward: [[78.29179]], Actor Loss: [[-4.423343e-02
-9.819363e-02 -5.294963e+01]], Critic Loss: 2818.766357421875
Episode 362/500, Total Reward: [[63.224102]], Actor Loss: [[-2.7359396e-02
-5.9705102e-01 -5.1982677e+01]], Critic Loss: 2767.505615234375
Episode 363/500, Total Reward: [[69.24181]], Actor Loss: [[-1.9108282e-02
-2.7185550e-01 -6.0687294e+01]], Critic Loss: 3718.34765625
Episode 364/500, Total Reward: [[80.839836]], Actor Loss: [[-8.6164819e-03
-1.9442747e+00 -3.8927860e+01]], Critic Loss: 1671.2359619140625
Episode 365/500, Total Reward: [[100.357]], Actor Loss: [[-7.6670356e-02
-5.6781858e-01 -9.1701950e+01]], Critic Loss: 8527.865234375
```

```
Episode 366/500, Total Reward: [[66.89463]], Actor Loss: [[-1.3386856e-03
-1.0032170e-02 -3.8780327e+01]], Critic Loss: 1504.796142578125
Episode 367/500, Total Reward: [[144.14453]], Actor Loss: [[-2.1588632e-03
-2.2204861e-02 -1.4435704e+02]], Critic Loss: 20845.990234375
Episode 368/500, Total Reward: [[93.23197]], Actor Loss: [[-3.8337451e-04
-1.2831987e-02 -1.0021480e+02]], Critic Loss: 10045.654296875
Episode 369/500, Total Reward: [[95.52905]], Actor Loss: [[-5.5625779e-03
-3.7459042e-02 -1.3546922e+02]], Critic Loss: 18363.56640625
Episode 370/500, Total Reward: [[133.02464]], Actor Loss: [[-1.33736583e-04
-9.98001266e-03 -1.06875885e+02]], Critic Loss: 11424.6171875
Episode 371/500, Total Reward: [[115.8227]], Actor Loss: [[-9.63838247e-05
-1.28383692e-02 -1.04045166e+02]], Critic Loss: 10828.0888671875
Episode 372/500, Total Reward: [[108.36879]], Actor Loss: [[-1.3560328e-05
-1.0743431e-03 -9.8817383e+01]], Critic Loss: 9765.0888671875
Episode 373/500, Total Reward: [[128.31206]], Actor Loss: [[-6.8214338e-04
-2.3650924e-02 -1.6298538e+02]], Critic Loss: 26572.1640625
Episode 374/500, Total Reward: [[90.66807]], Actor Loss: [[-1.0872878e-02
-8.8064775e-02 -7.4559288e+01]], Critic Loss: 5573.849609375
Episode 375/500, Total Reward: [[96.99178]], Actor Loss: [[ -0.21004465
-0.3897283  -87.99087   ]], Critic Loss: 7848.30078125
Episode 376/500, Total Reward: [[106.7974]], Actor Loss: [[  -0.41760087
-1.1161317  -110.68985   ]], Critic Loss: 12594.1337890625
Episode 377/500, Total Reward: [[94.96692]], Actor Loss: [[-1.1280456e-02
-8.4232770e-02 -6.8717941e+01]], Critic Loss: 4735.29248046875
Episode 378/500, Total Reward: [[110.652725]], Actor Loss: [[-3.946104e-04
-9.351328e-03 -8.796145e+01]], Critic Loss: 7738.9306640625
Episode 379/500, Total Reward: [[112.38292]], Actor Loss: [[-1.6232062e-04
-2.3349408e-02 -6.1390793e+01]], Critic Loss: 3771.717041015625
Episode 380/500, Total Reward: [[83.27297]], Actor Loss: [[-2.4784039e-04
-4.1291174e-03 -1.0837522e+02]], Critic Loss: 11746.1376953125
Episode 381/500, Total Reward: [[71.08904]], Actor Loss: [[-8.5311905e-05
-8.3756866e-04 -6.8551491e+01]], Critic Loss: 4699.43359375
Episode 382/500, Total Reward: [[75.60194]], Actor Loss: [[-8.3082006e-05
-1.6583410e-03 -9.0919838e+01]], Critic Loss: 8266.734375
Episode 383/500, Total Reward: [[129.15132]], Actor Loss: [[-8.3418740e-07
-4.2516485e-04 -9.2609360e+01]], Critic Loss: 8576.572265625
Episode 384/500, Total Reward: [[60.176563]], Actor Loss: [[-1.5632031e-05
-2.5561656e-04 -8.4961739e+01]], Critic Loss: 7218.54345703125
Episode 385/500, Total Reward: [[83.11059]], Actor Loss: [[-1.2042336e-05
-9.2767290e-04 -7.9545715e+01]], Critic Loss: 6327.669921875
Episode 386/500, Total Reward: [[121.6751]], Actor Loss: [[-2.706791e-05
-6.606515e-04 -1.383823e+02]], Critic Loss: 19149.849609375
Episode 387/500, Total Reward: [[141.92879]], Actor Loss: [[-1.2932377e-05
-9.4090437e-04 -1.4642592e+02]], Critic Loss: 21440.83203125
Episode 388/500, Total Reward: [[91.237816]], Actor Loss: [[-5.4446878e-06
-1.9238178e-04 -8.3960632e+01]], Critic Loss: 7049.4208984375
Episode 389/500, Total Reward: [[98.588684]], Actor Loss: [[-8.1103212e-07
-1.5564892e-04 -7.3652199e+01]], Critic Loss: 5424.669921875
```

```
Episode 390/500, Total Reward: [[141.5225]], Actor Loss: [[-7.8986021e-05
-1.5618524e-03 -1.9357350e+02]], Critic Loss: 37471.33203125
Episode 391/500, Total Reward: [[47.269985]], Actor Loss: [[-1.3436598e-04
-2.3584946e-03 -6.7489738e+01]], Critic Loss: 4555.20166015625
Episode 392/500, Total Reward: [[77.72265]], Actor Loss: [[-6.472070e-04
-7.457393e-03 -9.635972e+01]], Critic Loss: 9286.7587890625
Episode 393/500, Total Reward: [[107.10765]], Actor Loss: [[-1.6767026e-06
-1.4448416e-04 -7.9710289e+01]], Critic Loss: 6353.75341796875
Episode 394/500, Total Reward: [[78.781166]], Actor Loss: [[-5.3043270e-05
-1.1224351e-03 -1.0034571e+02]], Critic Loss: 10069.4970703125
Episode 395/500, Total Reward: [[112.232155]], Actor Loss: [[-5.5717346e-06
-9.6371712e-04 -9.1463028e+01]], Critic Loss: 8365.6630859375
Episode 396/500, Total Reward: [[72.60686]], Actor Loss: [[-2.1333241e-05
-2.1518498e-04 -7.8786095e+01]], Critic Loss: 6207.2861328125
Episode 397/500, Total Reward: [[97.93835]], Actor Loss: [[-4.2277372e-07
-9.4033530e-05 -7.6995514e+01]], Critic Loss: 5928.3232421875
Episode 398/500, Total Reward: [[76.121796]], Actor Loss: [[-4.0031600e-06
-1.6416798e-04 -7.6548103e+01]], Critic Loss: 5859.6376953125
Episode 399/500, Total Reward: [[44.507603]], Actor Loss: [[-2.6748348e-05
-5.2442617e-04 -4.9503479e+01]], Critic Loss: 2450.64892578125
Episode 400/500, Total Reward: [[187.93448]], Actor Loss: [[-3.5267865e-07
-1.1080146e-04 -1.4867628e+02]], Critic Loss: 22104.669921875
Episode 401/500, Total Reward: [[57.297436]], Actor Loss: [[-8.2611459e-06
-2.5634607e-04 -8.0773354e+01]], Critic Loss: 6524.37646484375
Episode 402/500, Total Reward: [[34.2824]], Actor Loss: [[-5.3175329e-04
-2.1209139e-03 -5.9310314e+01]], Critic Loss: 3518.02783203125
Episode 403/500, Total Reward: [[60.74684]], Actor Loss: [[-1.8288438e-05
-1.3373932e-03 -3.5273380e+01]], Critic Loss: 1244.306884765625
Episode 404/500, Total Reward: [[62.94714]], Actor Loss: [[-1.5891706e-04
-4.0954305e-03 -5.9621368e+01]], Critic Loss: 3555.21533203125
Episode 405/500, Total Reward: [[105.72224]], Actor Loss: [[-1.5778132e-05
-7.2092272e-04 -9.7682953e+01]], Critic Loss: 9542.1025390625
Episode 406/500, Total Reward: [[94.64857]], Actor Loss: [[-2.0618549e-04
-5.3026648e-03 -8.9220421e+01]], Critic Loss: 7961.2666015625
Episode 407/500, Total Reward: [[119.42495]], Actor Loss: [[-4.5266643e-04
-1.0306120e-02 -1.4523235e+02]], Critic Loss: 21095.55859375
Episode 408/500, Total Reward: [[132.53143]], Actor Loss: [[-3.6937502e-04
-6.3855853e-03 -1.2723985e+02]], Critic Loss: 16191.6982421875
Episode 409/500, Total Reward: [[94.11093]], Actor Loss: [[-2.1882076e-04
-7.7051837e-03 -7.5033875e+01]], Critic Loss: 5631.27197265625
Episode 410/500, Total Reward: [[110.819595]], Actor Loss: [[-8.2717231e-04
-6.6362722e-03 -1.0813604e+02]], Critic Loss: 11695.0166015625
Episode 411/500, Total Reward: [[114.81774]], Actor Loss: [[-3.10357194e-04
-4.40507615e-03 -1.15818825e+02]], Critic Loss: 13415.09375
Episode 412/500, Total Reward: [[121.308266]], Actor Loss: [[-1.49644975e-05
-1.35508960e-03 -1.37158569e+02]], Critic Loss: 18812.849609375
Episode 413/500, Total Reward: [[120.72953]], Actor Loss: [[-9.8897444e-06
-3.8695944e-04 -1.2179032e+02]], Critic Loss: 14832.9775390625
```

```
Episode 414/500, Total Reward: [[82.307594]], Actor Loss: [[-1.5271775e-03
-5.3368132e-03 -8.5994125e+01]], Critic Loss: 7396.17041015625
Episode 415/500, Total Reward: [[57.5742]], Actor Loss: [[-3.4144404e-04
-3.3544223e-03 -7.7387787e+01]], Critic Loss: 5989.4423828125
Episode 416/500, Total Reward: [[56.043922]], Actor Loss: [[-5.5552351e-05
-1.4391915e-03 -7.0236755e+01]], Critic Loss: 4933.412109375
Episode 417/500, Total Reward: [[141.84404]], Actor Loss: [[-1.3403808e-05
-4.3027365e-04 -1.4331772e+02]], Critic Loss: 20540.095703125
Episode 418/500, Total Reward: [[60.63321]], Actor Loss: [[-4.8548768e-06
-2.6706472e-04 -4.8959587e+01]], Critic Loss: 2397.068115234375
Episode 419/500, Total Reward: [[72.16632]], Actor Loss: [[-2.576132e-06
-9.063009e-05 -8.231316e+01]], Critic Loss: 6775.47216796875
Episode 420/500, Total Reward: [[63.032673]], Actor Loss: [[-4.7397248e-07
-1.9522291e-04 -7.7898872e+01]], Critic Loss: 6068.26513671875
Episode 421/500, Total Reward: [[44.01507]], Actor Loss: [[-1.2006375e-06
-9.3772942e-05 -5.7879047e+01]], Critic Loss: 3349.9951171875
Episode 422/500, Total Reward: [[53.20553]], Actor Loss: [[-1.1492331e-06
-3.3557470e-05 -5.4768719e+01]], Critic Loss: 2999.6162109375
Episode 423/500, Total Reward: [[73.97988]], Actor Loss: [[-1.5195319e-07
-2.7651511e-05 -6.2231041e+01]], Critic Loss: 3872.706298828125
Episode 424/500, Total Reward: [[93.05509]], Actor Loss: [[-1.9217923e-07
-3.9037015e-05 -1.1447748e+02]], Critic Loss: 13105.1015625
Episode 425/500, Total Reward: [[67.89516]], Actor Loss: [[-1.2781994e-07
-3.5363071e-05 -8.9140404e+01]], Critic Loss: 7946.01708984375
Episode 426/500, Total Reward: [[71.20686]], Actor Loss: [[-6.4349712e-07
-1.3968612e-04 -8.6325203e+01]], Critic Loss: 7452.06591796875
Episode 427/500, Total Reward: [[79.68221]], Actor Loss: [[-1.8797697e-07
-1.3574135e-05 -8.5508331e+01]], Critic Loss: 7311.67626953125
Episode 428/500, Total Reward: [[67.32864]], Actor Loss: [[-2.631850e-06
-9.657055e-05 -6.830275e+01]], Critic Loss: 4665.279296875
Episode 429/500, Total Reward: [[77.25089]], Actor Loss: [[-1.0431403e-06
-6.2157371e-05 -8.8226768e+01]], Critic Loss: 7783.9736328125
Episode 430/500, Total Reward: [[154.09373]], Actor Loss: [[-5.4064782e-07
-6.5125030e-05 -1.5762537e+02]], Critic Loss: 24845.775390625
Episode 431/500, Total Reward: [[118.35907]], Actor Loss: [[-1.3164828e-05
-2.4089069e-04 -1.4013348e+02]], Critic Loss: 19637.4609375
Episode 432/500, Total Reward: [[95.719345]], Actor Loss: [[-9.7918339e-05
-1.1194776e-03 -1.0131617e+02]], Critic Loss: 10265.2138671875
Episode 433/500, Total Reward: [[73.887535]], Actor Loss: [[-3.6287565e-05
-6.2443677e-04 -7.5644646e+01]], Critic Loss: 5722.2119140625
Episode 434/500, Total Reward: [[66.995415]], Actor Loss: [[-6.9131126e-04
-3.3226828e-03 -6.4649078e+01]], Critic Loss: 4180.0224609375
Episode 435/500, Total Reward: [[108.52364]], Actor Loss: [[-4.7030585e-04
-7.8876466e-03 -1.1160705e+02]], Critic Loss: 12457.998046875
Episode 436/500, Total Reward: [[105.83247]], Actor Loss: [[-4.8632347e-03
-1.3844820e-02 -1.0965900e+02]], Critic Loss: 12029.1982421875
Episode 437/500, Total Reward: [[76.62762]], Actor Loss: [[-6.3944125e-04
-1.5067325e-02 -8.3969704e+01]], Critic Loss: 7053.54833984375
```

```
Episode 438/500, Total Reward: [[92.095856]], Actor Loss: [[-1.3802082e-03
-1.0285194e-01 -8.6404305e+01]], Critic Loss: 7483.72705078125
Episode 439/500, Total Reward: [[62.118637]], Actor Loss: [[-4.9725384e-03
-3.7733689e-02 -5.0052105e+01]], Critic Loss: 2509.489990234375
Episode 440/500, Total Reward: [[83.79487]], Actor Loss: [[-4.1230605e-03
-8.0457062e-02 -7.2676514e+01]], Critic Loss: 5294.1767578125
Episode 441/500, Total Reward: [[87.47881]], Actor Loss: [[-5.6911223e-03
-1.2995429e-02 -7.7708130e+01]], Critic Loss: 6041.45751953125
Episode 442/500, Total Reward: [[48.46181]], Actor Loss: [[-7.2984085e-03
-1.9838082e-02 -4.9880787e+01]], Critic Loss: 2490.801025390625
Episode 443/500, Total Reward: [[107.56332]], Actor Loss: [[-1.50229009e-02
-4.92292196e-02 -1.23782265e+02]], Critic Loss: 15337.9580078125
Episode 444/500, Total Reward: [[172.31548]], Actor Loss: [[  -0.23815496
-1.0059935  -172.61452   ]], Critic Loss: 30226.83203125
Episode 445/500, Total Reward: [[66.27143]], Actor Loss: [[ -2.6406431
-1.7131859 -82.723    ]], Critic Loss: 7582.3740234375
Episode 446/500, Total Reward: [[82.361626]], Actor Loss: [[ -1.6854367
-1.1744082 -80.44294  ]], Critic Loss: 6939.35302734375
Episode 447/500, Total Reward: [[66.92578]], Actor Loss: [[-15.735746
-6.3951783 -37.133034 ]], Critic Loss: 3512.21630859375
Episode 448/500, Total Reward: [[131.03252]], Actor Loss: [[-48.161392 -40.7904
-57.36667 ]], Critic Loss: 21409.09375
Episode 449/500, Total Reward: [[119.98501]], Actor Loss: [[-56.61222
-19.076939 -46.445797]], Critic Loss: 14916.947265625
Episode 450/500, Total Reward: [[92.07882]], Actor Loss: [[-26.090033
-35.020832  -15.6477585]], Critic Loss: 5891.8857421875
Episode 451/500, Total Reward: [[74.80711]], Actor Loss: [[-44.43534  -12.379146
-30.981394]], Critic Loss: 7708.11572265625
Episode 452/500, Total Reward: [[124.53235]], Actor Loss: [[-12.653412 -93.60472
-15.34223 ]], Critic Loss: 14786.646484375
Episode 453/500, Total Reward: [[65.80009]], Actor Loss: [[-40.722828
-7.6094294 -16.682423 ]], Critic Loss: 4226.90869140625
Episode 454/500, Total Reward: [[114.26461]], Actor Loss: [[-66.136536
-9.5268345 -32.38619  ]], Critic Loss: 11674.7060546875
Episode 455/500, Total Reward: [[87.82198]], Actor Loss: [[-31.497562 -16.289751
-44.111492]], Critic Loss: 8445.3896484375
Episode 456/500, Total Reward: [[40.569046]], Actor Loss: [[-20.156235
-9.7323065 -25.570297 ]], Critic Loss: 3075.682373046875
Episode 457/500, Total Reward: [[91.43821]], Actor Loss: [[-27.752665 -30.122646
-26.718458]], Critic Loss: 7156.1064453125
Episode 458/500, Total Reward: [[113.74893]], Actor Loss: [[-14.8756695
-70.47705   -10.246437 ]], Critic Loss: 9139.1982421875
Episode 459/500, Total Reward: [[97.891426]], Actor Loss: [[-60.93712
-3.8834872 -30.786863 ]], Critic Loss: 9140.7880859375
Episode 460/500, Total Reward: [[111.02523]], Actor Loss: [[-64.18639
-8.424563 -47.347534]], Critic Loss: 14390.0390625
Episode 461/500, Total Reward: [[87.63421]], Actor Loss: [[-39.276394  -9.279972
-36.200157]], Critic Loss: 7183.66943359375
```

Episode 462/500, Total Reward: [[99.15609]], Actor Loss: [[-45.892242 -4.548925
-40.461887]], Critic Loss: 8263.365234375
Episode 463/500, Total Reward: [[137.43642]], Actor Loss: [[-83.253136
-4.9265738 -40.560566]], Critic Loss: 16574.056640625
Episode 464/500, Total Reward: [[65.50659]], Actor Loss: [[-20.48895   -9.212117
-28.740944]], Critic Loss: 3415.46923828125
Episode 465/500, Total Reward: [[56.24589]], Actor Loss: [[-11.206952
-4.1396265 -33.77877]], Critic Loss: 2413.300048828125
Episode 466/500, Total Reward: [[60.03612]], Actor Loss: [[-9.575652 -4.670888
-56.95089]], Critic Loss: 5069.07470703125
Episode 467/500, Total Reward: [[64.0196]], Actor Loss: [[-10.796345
-3.9469569 -51.816425]], Critic Loss: 4430.19677734375
Episode 468/500, Total Reward: [[51.183796]], Actor Loss: [[-7.6110983
-3.7106442 -57.07349]], Critic Loss: 4677.90771484375
Episode 469/500, Total Reward: [[57.744595]], Actor Loss: [[-11.925885
-3.302666 -26.262545]], Critic Loss: 1721.5111083984375
Episode 470/500, Total Reward: [[53.66353]], Actor Loss: [[-13.244548 -4.621926
-60.747498]], Critic Loss: 6180.15576171875
Episode 471/500, Total Reward: [[100.13291]], Actor Loss: [[-14.891125
-5.2211237 -92.04781]], Critic Loss: 12579.8798828125
Episode 472/500, Total Reward: [[209.73193]], Actor Loss: [[-29.685516
-17.132816 -157.72348]], Critic Loss: 41837.3515625
Episode 473/500, Total Reward: [[59.16987]], Actor Loss: [[-15.039646 -5.567515
-32.77485]], Critic Loss: 2849.63916015625
Episode 474/500, Total Reward: [[71.49659]], Actor Loss: [[-17.506079
-4.3012943 -46.34561]], Critic Loss: 4644.8291015625
Episode 475/500, Total Reward: [[47.269985]], Actor Loss: [[-7.893515
-2.9602375 -23.756447]], Critic Loss: 1197.8658447265625
Episode 476/500, Total Reward: [[62.563198]], Actor Loss: [[-8.9034
-6.0075545 -46.741905]], Critic Loss: 3801.075439453125
Episode 477/500, Total Reward: [[61.787018]], Actor Loss: [[-12.206636
-3.4619455 -52.793182]], Critic Loss: 4687.0126953125
Episode 478/500, Total Reward: [[80.90446]], Actor Loss: [[-12.432982
-4.9495935 -67.798935]], Critic Loss: 7255.888671875
Episode 479/500, Total Reward: [[101.74885]], Actor Loss: [[-47.05259
-17.644958 -36.242203]], Critic Loss: 10188.8330078125
Episode 480/500, Total Reward: [[70.678154]], Actor Loss: [[-43.0068
-11.422522 -15.854731]], Critic Loss: 4939.8486328125
Episode 481/500, Total Reward: [[156.1771]], Actor Loss: [[-110.827736
-17.558252 -35.064117]], Critic Loss: 26715.935546875
Episode 482/500, Total Reward: [[82.148445]], Actor Loss: [[-74.92217
-7.4451613 -19.14967]], Critic Loss: 10305.701171875
Episode 483/500, Total Reward: [[70.244354]], Actor Loss: [[-56.258232
-5.985971 -16.556288]], Critic Loss: 6209.517578125
Episode 484/500, Total Reward: [[58.199894]], Actor Loss: [[-46.50025
-7.738653 -8.055311]], Critic Loss: 3880.56982421875
Episode 485/500, Total Reward: [[58.91773]], Actor Loss: [[-45.519695 -5.774748
-12.31898]], Critic Loss: 4046.66748046875

```
Episode 486/500, Total Reward: [[65.71092]], Actor Loss: [[-41.75299   -9.152448
-8.811363]], Critic Loss: 3566.0966796875
Episode 487/500, Total Reward: [[47.933292]], Actor Loss: [[-43.11328
-1.9986546  -5.4555287]], Critic Loss: 2557.068359375
Episode 488/500, Total Reward: [[138.29964]], Actor Loss: [[-118.1586
-6.102407   -2.887437]], Critic Loss: 16166.7255859375
Episode 489/500, Total Reward: [[113.33506]], Actor Loss: [[-62.5027
-20.541052 -34.49124 ]], Critic Loss: 13814.4736328125
Episode 490/500, Total Reward: [[76.68203]], Actor Loss: [[-34.429592 -25.026258
-26.082378]], Critic Loss: 7316.78759765625
Episode 491/500, Total Reward: [[144.01979]], Actor Loss: [[-66.02667
-32.700104 -39.981873]], Critic Loss: 19240.08984375
Episode 492/500, Total Reward: [[93.677864]], Actor Loss: [[-36.01357
-22.833841 -31.175829]], Critic Loss: 8104.18505859375
Episode 493/500, Total Reward: [[60.82867]], Actor Loss: [[-30.900667 -11.102884
-16.951008]], Critic Loss: 3475.64013671875
Episode 494/500, Total Reward: [[95.21897]], Actor Loss: [[-35.102173 -20.590229
-41.51795 ]], Critic Loss: 9449.8525390625
Episode 495/500, Total Reward: [[61.925]], Actor Loss: [[-21.698236 -10.46524
-19.692455]], Critic Loss: 2689.03759765625
Episode 496/500, Total Reward: [[79.446526]], Actor Loss: [[-22.490875
-18.870075 -37.93217 ]], Critic Loss: 6287.400390625
Episode 497/500, Total Reward: [[101.06904]], Actor Loss: [[-22.34889
-16.298363 -62.030125]], Critic Loss: 10135.93359375
Episode 498/500, Total Reward: [[130.70998]], Actor Loss: [[-74.53623
-14.929911 -43.421696]], Critic Loss: 17659.17578125
Episode 499/500, Total Reward: [[73.83059]], Actor Loss: [[-47.38876  -11.313825
-20.339943]], Critic Loss: 6247.72119140625
Episode 500/500, Total Reward: [[108.49406]], Actor Loss: [[-59.204086
-19.288769 -24.209415]], Critic Loss: 10547.7568359375
```

Evaluate the performance of the model on test set (0.5 M)

[ ]:
```python
def evaluate_model(actor_model, energy_model, X_test):
    """
    Evaluate the trained Actor-Critic model on the test set.

    Parameters:
    - actor_model: Trained Actor model
    - energy_model: Trained energy prediction model
    - X_test: Testing state features (numpy array)

    Returns:
    - total_reward: Cumulative reward over the test set (float)
    """
    total_reward = 0

    for state in X_test:
```

```
        # Get action probabilities from the Actor
        action_probs = actor_model.predict(state.reshape(1, -1), verbose=0)
        action = np.random.choice(action_space, p=action_probs.flatten())  #␣
↪Sample action based on probabilities

        # Simulate environment: get next_state, y_next, and reward
        next_state, y_next, reward = simulate_environment(state, action,␣
↪energy_model)

        # Accumulate reward
        total_reward += reward

    print(f"Total reward on the test set: {total_reward:.2f}")
```

[ ]:
```
# Evaluate the model on the test set
evaluate_model(actor_model, energy_model, X_test)
```

```
Streaming output truncated to the last 5000 lines.
1/1               0s 24ms/step
1/1               0s 19ms/step
1/1               0s 25ms/step
1/1               0s 19ms/step
1/1               0s 21ms/step
1/1               0s 26ms/step
1/1               0s 25ms/step
1/1               0s 26ms/step
1/1               0s 29ms/step
1/1               0s 29ms/step
1/1               0s 34ms/step
1/1               0s 24ms/step
1/1               0s 59ms/step
1/1               0s 28ms/step
1/1               0s 42ms/step
1/1               0s 35ms/step
1/1               0s 42ms/step
1/1               0s 33ms/step
1/1               0s 36ms/step
1/1               0s 30ms/step
1/1               0s 28ms/step
1/1               0s 38ms/step
1/1               0s 35ms/step
1/1               0s 41ms/step
1/1               0s 33ms/step
1/1               0s 43ms/step
1/1               0s 64ms/step
1/1               0s 18ms/step
1/1               0s 24ms/step
```

```
1/1                0s 18ms/step
1/1                0s 26ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 32ms/step
1/1                0s 18ms/step
1/1                0s 20ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 23ms/step
1/1                0s 19ms/step
1/1                0s 28ms/step
1/1                0s 18ms/step
1/1                0s 29ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 20ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 18ms/step
1/1                0s 22ms/step
1/1                0s 27ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 19ms/step
1/1                0s 23ms/step
1/1                0s 29ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 21ms/step
1/1                0s 21ms/step
1/1                0s 18ms/step
1/1                0s 18ms/step
1/1                0s 27ms/step
1/1                0s 17ms/step
1/1                0s 27ms/step
```

```
1/1                0s 17ms/step
1/1                0s 20ms/step
1/1                0s 22ms/step
1/1                0s 19ms/step
1/1                0s 18ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 22ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 19ms/step
1/1                0s 18ms/step
1/1                0s 17ms/step
1/1                0s 21ms/step
1/1                0s 18ms/step
1/1                0s 27ms/step
1/1                0s 19ms/step
1/1                0s 21ms/step
1/1                0s 19ms/step
1/1                0s 19ms/step
1/1                0s 17ms/step
1/1                0s 27ms/step
1/1                0s 26ms/step
1/1                0s 22ms/step
1/1                0s 29ms/step
1/1                0s 31ms/step
1/1                0s 32ms/step
1/1                0s 30ms/step
1/1                0s 37ms/step
1/1                0s 32ms/step
1/1                0s 41ms/step
1/1                0s 36ms/step
1/1                0s 25ms/step
1/1                0s 31ms/step
1/1                0s 43ms/step
1/1                0s 48ms/step
1/1                0s 25ms/step
1/1                0s 28ms/step
```

```
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 49ms/step
1/1                0s 42ms/step
1/1                0s 45ms/step
1/1                0s 47ms/step
1/1                0s 18ms/step
1/1                0s 23ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 26ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 31ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 19ms/step
1/1                0s 22ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
```

```
1/1                0s 16ms/step
1/1                0s 23ms/step
1/1                0s 19ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 20ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 23ms/step
1/1                0s 21ms/step
1/1                0s 28ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 30ms/step
1/1                0s 16ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 20ms/step
1/1                0s 19ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 19ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 19ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 28ms/step
1/1                0s 27ms/step
1/1                0s 29ms/step
```

```
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 41ms/step
1/1                0s 23ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 26ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 40ms/step
1/1                0s 23ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 29ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 28ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 34ms/step
1/1                0s 31ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 30ms/step
1/1                0s 24ms/step
1/1                0s 31ms/step
1/1                0s 25ms/step
1/1                0s 27ms/step
1/1                0s 30ms/step
1/1                0s 38ms/step
1/1                0s 39ms/step
1/1                0s 35ms/step
1/1                0s 27ms/step
1/1                0s 26ms/step
1/1                0s 27ms/step
```

```
1/1              0s 32ms/step
1/1              0s 35ms/step
1/1              0s 28ms/step
1/1              0s 36ms/step
1/1              0s 48ms/step
1/1              0s 41ms/step
1/1              0s 44ms/step
1/1              0s 50ms/step
1/1              0s 30ms/step
1/1              0s 28ms/step
1/1              0s 25ms/step
1/1              0s 31ms/step
1/1              0s 25ms/step
1/1              0s 59ms/step
1/1              0s 35ms/step
1/1              0s 35ms/step
1/1              0s 35ms/step
1/1              0s 35ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 35ms/step
1/1              0s 23ms/step
1/1              0s 26ms/step
1/1              0s 25ms/step
1/1              0s 33ms/step
1/1              0s 25ms/step
1/1              0s 27ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 25ms/step
1/1              0s 29ms/step
1/1              0s 31ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 23ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 27ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 23ms/step
1/1              0s 29ms/step
1/1              0s 30ms/step
1/1              0s 24ms/step
1/1              0s 25ms/step
```

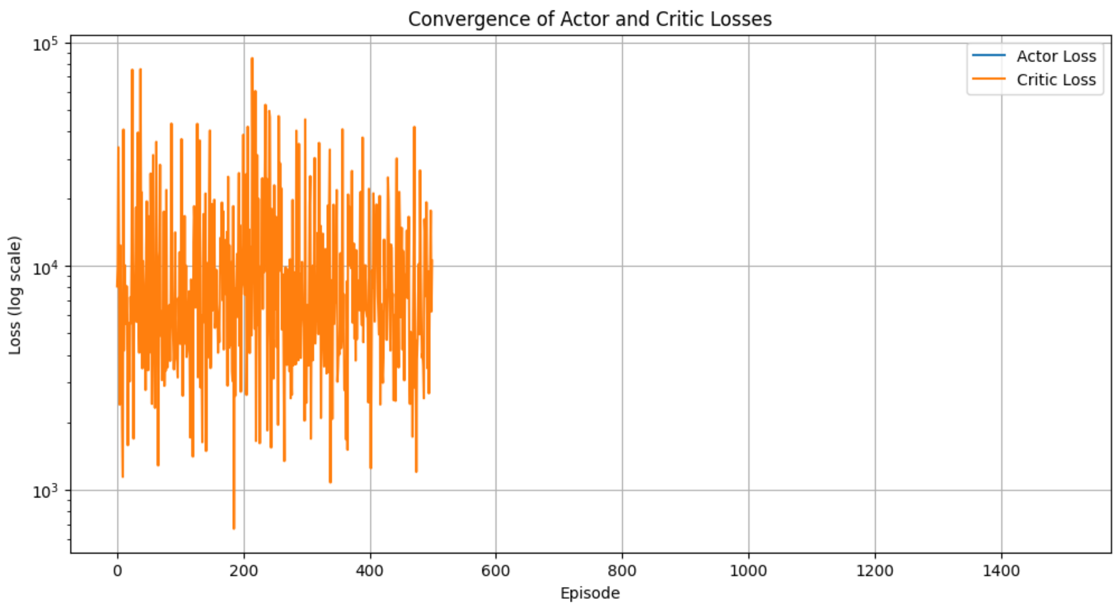
```
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 32ms/step
1/1                0s 26ms/step
1/1                0s 38ms/step
1/1                0s 23ms/step
1/1                0s 23ms/step
1/1                0s 22ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 39ms/step
1/1                0s 23ms/step
1/1                0s 26ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 26ms/step
1/1                0s 27ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 38ms/step
1/1                0s 23ms/step
1/1                0s 23ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 28ms/step
1/1                0s 30ms/step
1/1                0s 23ms/step
1/1                0s 31ms/step
1/1                0s 23ms/step
1/1                0s 34ms/step
1/1                0s 30ms/step
1/1                0s 26ms/step
```

```
1/1              0s 24ms/step
1/1              0s 25ms/step
1/1              0s 23ms/step
1/1              0s 23ms/step
1/1              0s 30ms/step
1/1              0s 26ms/step
1/1              0s 49ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 23ms/step
1/1              0s 25ms/step
1/1              0s 30ms/step
1/1              0s 36ms/step
1/1              0s 41ms/step
1/1              0s 50ms/step
1/1              0s 33ms/step
1/1              0s 53ms/step
1/1              0s 39ms/step
1/1              0s 37ms/step
1/1              0s 25ms/step
1/1              0s 44ms/step
1/1              0s 37ms/step
1/1              0s 50ms/step
1/1              0s 32ms/step
1/1              0s 30ms/step
1/1              0s 31ms/step
Total reward on the test set: -124187.53
```

1.0.2 Plot the convergence of Actor and Critic losses (1 M)

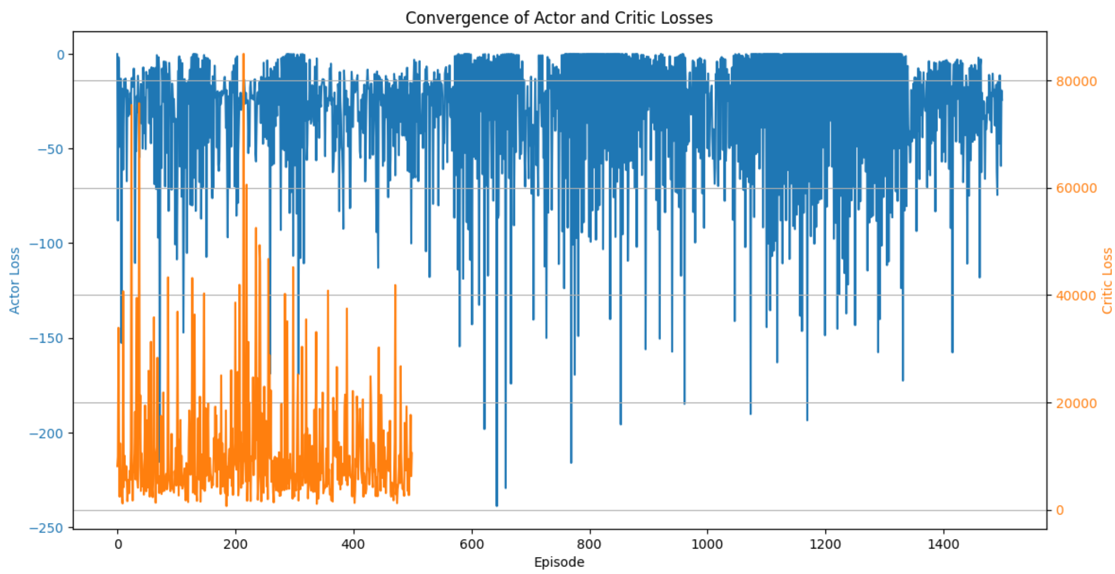
[ ]:
```python
import numpy as np
import matplotlib.pyplot as plt
def plot_convergence(actor_losses, critic_losses):
    """
    Plot the convergence of Actor and Critic losses over episodes with log␣
 ↪scale.
    """
    actor_losses = np.array(actor_losses).reshape(-1)
    critic_losses = np.array(critic_losses).reshape(-1)

    plt.figure(figsize=(12, 6))
    plt.plot(actor_losses, label='Actor Loss')
    plt.plot(critic_losses, label='Critic Loss')
    plt.yscale('log')  # Set y-axis to log scale
    plt.xlabel('Episode')
    plt.ylabel('Loss (log scale)')
    plt.title('Convergence of Actor and Critic Losses')
```

```
    plt.legend()
    plt.grid(True)
    plt.show()

# Call the function to plot
plot_convergence(actor_losses, critic_losses)
```



[ ]:
```
import numpy as np
import matplotlib.pyplot as plt
def plot_convergence(actor_losses, critic_losses):
    """
    Plot the convergence of Actor and Critic losses over episodes using two␣
 ↪y-axes.
    """
    actor_losses = np.array(actor_losses).reshape(-1)
    critic_losses = np.array(critic_losses).reshape(-1)

    fig, ax1 = plt.subplots(figsize=(12, 6))

    color = 'tab:blue'
    ax1.set_xlabel('Episode')
    ax1.set_ylabel('Actor Loss', color=color)
    ax1.plot(actor_losses, color=color, label='Actor Loss')
    ax1.tick_params(axis='y', labelcolor=color)

    ax2 = ax1.twinx()  # Instantiate a second axes that shares the same x-axis
    color = 'tab:orange'
```

```
    ax2.set_ylabel('Critic Loss', color=color)  # We already handled the␣
↪x-label with ax1
    ax2.plot(critic_losses, color=color, label='Critic Loss')
    ax2.tick_params(axis='y', labelcolor=color)

    fig.tight_layout()  # Ensure everything fits without overlap
    plt.title('Convergence of Actor and Critic Losses')
    plt.grid(True)
    plt.show()

# Call the function to plot
plot_convergence(actor_losses, critic_losses)
```



1.0.3 Plot the learned policy - by showing the action probabilities across different state values (1 M)

[ ]:
```
class YourEnvironment:
    def __init__(self, target_temp=22):
        self.state = np.zeros(22)  # Adjust state vector to have 22 features
        self.done = False
        self.action_space = 3  # Example: 0=Decrease, 1=Maintain, 2=Increase
        self.target_temp = target_temp  # Define target_temp attribute

    def reset(self):
        self.state = np.zeros(22)  # Reset state to initial values
        self.done = False
        return self.state
```

```
    def step(self, action):
        temp_adjustment = [-1, 0, 1]  # Decrease, Maintain, Increase
        next_state = self.state.copy()
        next_state[:9] += temp_adjustment[action]  # Apply action to␣
↪temperature features

        # Simulate energy consumption (replace with actual model prediction)
        y_next = np.random.random()  # Placeholder for energy model prediction
        y_current = np.random.random()  # Placeholder for current energy␣
↪consumption

        # Calculate reward based on current and next state
        reward = self.calculate_reward(self.state, next_state, y_current,␣
↪y_next)

        self.state = next_state
        self.done = np.random.rand() > 0.95  # Randomly set done for this␣
↪example

        return next_state, reward, self.done, {}

    def render(self):
        # Optional: Display or print the current state (for visualization/
↪debugging)
        print(f"Current State: {self.state}")

    def calculate_reward(self, state, next_state, y_current, y_next):
        # Define how the reward is calculated based on the state changes and␣
↪energy consumption
        temp_diff = np.abs(next_state[:9] - self.target_temp).mean()
        energy_saving = y_current - y_next
        reward = -temp_diff + energy_saving  # Example reward calculation
        return reward
```

[ ]:
```
def plot_learned_policy(actor_model, env):
    state = env.reset()
    done = False

    while not done:
        state_input = np.array(state).reshape(1, -1)
        action_probs = actor_model(state_input)
        action = np.argmax(action_probs)

        next_state, reward, done, _ = env.step(action)
        env.render()
```

```
        state = next_state

# Initialize your environment
env = YourEnvironment(target_temp=89)  # Replace with your actual environment␣
 ↪initialization

# # Train your actor-critic model
# actor_losses, critic_losses = train_actor_critic(
#     actor_model, critic_model, energy_model, X_train,
#     episodes=500, gamma=0.99
# )

# After training, call the plot_learned_policy to visualize the policy learned␣
 ↪by the actor
plot_learned_policy(actor_model, env)
```

```
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
Current State: [-1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  0.  0.  0.  0.  0.
0.  0.
  0.  0.  0.  0.]
```

Conclusion (0.5 M)

[ ]:
```
# Baseline average energy consumption
import numpy as np
baseline_energy = np.mean(y_test)
print(f"Baseline Average Energy Consumption: {baseline_energy:.2f} Wh")
```

```
Baseline Average Energy Consumption: 99.34 Wh
```

[ ]:
```
def calculate_energy_with_rl(actor_model, energy_model, X_test):
    """
    Calculate the average energy consumption on the test set with RL␣
↪adjustments.

    Parameters:
    - actor_model: Trained Actor model
    - energy_model: Trained energy prediction model
    - X_test: Testing state features (numpy array)

    Returns:
    - avg_energy_rl: Average energy consumption with RL (float)
    """
    total_energy = 0

    for state in X_test:
        # Get action probabilities from the Actor
        action_probs = actor_model.predict(state.reshape(1, -1), verbose=0)
        action = np.random.choice(action_space, p=action_probs.flatten())  #␣
↪Sample action based on probabilities

        # Simulate environment: get next_state, y_next, and reward
        next_state, y_next, _ = simulate_environment(state, action,␣
↪energy_model)

        # Accumulate energy consumption
        total_energy += y_next

    avg_energy_rl = total_energy / len(X_test)
    return avg_energy_rl

# Calculate average energy with RL
avg_energy_rl = calculate_energy_with_rl(actor_model, energy_model, X_test)
print(f"Average Energy Consumption with RL: {avg_energy_rl:.2f} Wh")
```

```
Streaming output truncated to the last 5000 lines.
1/1               0s 25ms/step
1/1               0s 18ms/step
1/1               0s 26ms/step
```

```
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 22ms/step
1/1                0s 27ms/step
1/1                0s 16ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
1/1                0s 22ms/step
1/1                0s 23ms/step
1/1                0s 19ms/step
1/1                0s 29ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
1/1                0s 22ms/step
1/1                0s 28ms/step
1/1                0s 16ms/step
1/1                0s 28ms/step
1/1                0s 18ms/step
1/1                0s 31ms/step
1/1                0s 18ms/step
1/1                0s 22ms/step
1/1                0s 16ms/step
1/1                0s 22ms/step
1/1                0s 20ms/step
1/1                0s 22ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 22ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 33ms/step
1/1                0s 31ms/step
1/1                0s 28ms/step
1/1                0s 33ms/step
1/1                0s 31ms/step
1/1                0s 37ms/step
1/1                0s 44ms/step
1/1                0s 35ms/step
1/1                0s 36ms/step
1/1                0s 33ms/step
```

```
1/1                0s 29ms/step
1/1                0s 29ms/step
1/1                0s 44ms/step
1/1                0s 48ms/step
1/1                0s 26ms/step
1/1                0s 34ms/step
1/1                0s 30ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 41ms/step
1/1                0s 43ms/step
1/1                0s 25ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 27ms/step
1/1                0s 18ms/step
1/1                0s 18ms/step
1/1                0s 21ms/step
1/1                0s 26ms/step
1/1                0s 19ms/step
1/1                0s 19ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 29ms/step
1/1                0s 16ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 22ms/step
```

```
1/1           0s 17ms/step
1/1           0s 23ms/step
1/1           0s 18ms/step
1/1           0s 25ms/step
1/1           0s 21ms/step
1/1           0s 23ms/step
1/1           0s 16ms/step
1/1           0s 24ms/step
1/1           0s 17ms/step
1/1           0s 24ms/step
1/1           0s 17ms/step
1/1           0s 22ms/step
1/1           0s 16ms/step
1/1           0s 24ms/step
1/1           0s 16ms/step
1/1           0s 21ms/step
1/1           0s 17ms/step
1/1           0s 25ms/step
1/1           0s 18ms/step
1/1           0s 23ms/step
1/1           0s 20ms/step
1/1           0s 22ms/step
1/1           0s 20ms/step
1/1           0s 27ms/step
1/1           0s 17ms/step
1/1           0s 23ms/step
1/1           0s 17ms/step
1/1           0s 25ms/step
1/1           0s 16ms/step
1/1           0s 23ms/step
1/1           0s 16ms/step
1/1           0s 29ms/step
1/1           0s 20ms/step
1/1           0s 24ms/step
1/1           0s 17ms/step
1/1           0s 24ms/step
1/1           0s 16ms/step
1/1           0s 25ms/step
1/1           0s 16ms/step
1/1           0s 26ms/step
1/1           0s 19ms/step
1/1           0s 28ms/step
1/1           0s 16ms/step
1/1           0s 22ms/step
1/1           0s 17ms/step
1/1           0s 24ms/step
1/1           0s 17ms/step
1/1           0s 24ms/step
```

```
1/1              0s 17ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 23ms/step
1/1              0s 17ms/step
1/1              0s 23ms/step
1/1              0s 17ms/step
1/1              0s 21ms/step
1/1              0s 17ms/step
1/1              0s 24ms/step
1/1              0s 26ms/step
1/1              0s 43ms/step
1/1              0s 27ms/step
1/1              0s 41ms/step
1/1              0s 24ms/step
1/1              0s 23ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 28ms/step
1/1              0s 46ms/step
1/1              0s 41ms/step
1/1              0s 25ms/step
1/1              0s 29ms/step
1/1              0s 31ms/step
1/1              0s 33ms/step
1/1              0s 29ms/step
1/1              0s 37ms/step
1/1              0s 35ms/step
1/1              0s 28ms/step
1/1              0s 39ms/step
1/1              0s 26ms/step
1/1              0s 36ms/step
1/1              0s 29ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 33ms/step
1/1              0s 24ms/step
1/1              0s 27ms/step
1/1              0s 33ms/step
1/1              0s 32ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 22ms/step
1/1              0s 24ms/step
1/1              0s 17ms/step
1/1              0s 20ms/step
1/1              0s 23ms/step
1/1              0s 23ms/step
```

```
1/1                0s 20ms/step
1/1                0s 22ms/step
1/1                0s 19ms/step
1/1                0s 25ms/step
1/1                0s 17ms/step
1/1                0s 26ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 18ms/step
1/1                0s 32ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 18ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 33ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 16ms/step
1/1                0s 28ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 22ms/step
1/1                0s 30ms/step
1/1                0s 17ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 24ms/step
1/1                0s 17ms/step
1/1                0s 25ms/step
1/1                0s 19ms/step
1/1                0s 24ms/step
1/1                0s 16ms/step
1/1                0s 24ms/step
1/1                0s 18ms/step
1/1                0s 23ms/step
1/1                0s 17ms/step
1/1                0s 21ms/step
1/1                0s 21ms/step
1/1                0s 21ms/step
1/1                0s 18ms/step
1/1                0s 22ms/step
1/1                0s 17ms/step
1/1                0s 32ms/step
```

```
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 27ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 27ms/step
1/1                0s 43ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 27ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 28ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 28ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 30ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 32ms/step
1/1                0s 47ms/step
1/1                0s 29ms/step
1/1                0s 47ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 28ms/step
1/1                0s 32ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 30ms/step
```

```
1/1                0s 35ms/step
1/1                0s 50ms/step
1/1                0s 44ms/step
1/1                0s 40ms/step
1/1                0s 26ms/step
1/1                0s 38ms/step
1/1                0s 28ms/step
1/1                0s 38ms/step
1/1                0s 49ms/step
1/1                0s 42ms/step
1/1                0s 38ms/step
1/1                0s 41ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 31ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 31ms/step
1/1                0s 28ms/step
1/1                0s 25ms/step
1/1                0s 30ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 30ms/step
1/1                0s 41ms/step
1/1                0s 28ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
```

```
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 22ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 33ms/step
1/1                0s 42ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 32ms/step
1/1                0s 45ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 34ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 29ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 46ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 29ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 37ms/step
1/1                0s 28ms/step
1/1                0s 41ms/step
```

```
1/1                0s 41ms/step
1/1                0s 53ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 46ms/step
1/1                0s 47ms/step
1/1                0s 50ms/step
1/1                0s 38ms/step
1/1                0s 42ms/step
1/1                0s 38ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 32ms/step
1/1                0s 44ms/step
1/1                0s 33ms/step
1/1                0s 29ms/step
1/1                0s 54ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 50ms/step
1/1                0s 25ms/step
1/1                0s 34ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 31ms/step
1/1                0s 35ms/step
1/1                0s 28ms/step
1/1                0s 33ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 29ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 41ms/step
1/1                0s 26ms/step
1/1                0s 29ms/step
1/1                0s 27ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 28ms/step
1/1                0s 27ms/step
1/1                0s 31ms/step
```

```
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 36ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 28ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 31ms/step
1/1                0s 36ms/step
1/1                0s 29ms/step
1/1                0s 24ms/step
1/1                0s 35ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 36ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 31ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 45ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 27ms/step
1/1                0s 35ms/step
1/1                0s 24ms/step
```

```
1/1                0s 25ms/step
1/1                0s 30ms/step
1/1                0s 29ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 40ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 37ms/step
1/1                0s 36ms/step
1/1                0s 45ms/step
1/1                0s 25ms/step
1/1                0s 39ms/step
1/1                0s 39ms/step
1/1                0s 28ms/step
1/1                0s 43ms/step
1/1                0s 32ms/step
1/1                0s 35ms/step
1/1                0s 23ms/step
1/1                0s 30ms/step
1/1                0s 29ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 39ms/step
1/1                0s 34ms/step
1/1                0s 45ms/step
1/1                0s 49ms/step
1/1                0s 32ms/step
1/1                0s 27ms/step
1/1                0s 34ms/step
1/1                0s 31ms/step
1/1                0s 29ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 34ms/step
1/1                0s 26ms/step
1/1                0s 30ms/step
1/1                0s 41ms/step
1/1                0s 24ms/step
1/1                0s 26ms/step
1/1                0s 43ms/step
1/1                0s 32ms/step
```

```
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 23ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 23ms/step
1/1                0s 23ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 25ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 29ms/step
1/1                0s 26ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 24ms/step
1/1                0s 32ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 26ms/step
1/1                0s 29ms/step
1/1                0s 26ms/step
1/1                0s 26ms/step
1/1                0s 25ms/step
1/1                0s 39ms/step
1/1                0s 25ms/step
1/1                0s 35ms/step
1/1                0s 25ms/step
1/1                0s 25ms/step
1/1                0s 29ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 33ms/step
1/1                0s 28ms/step
1/1                0s 34ms/step
1/1                0s 25ms/step
1/1                0s 24ms/step
1/1                0s 27ms/step
1/1                0s 25ms/step
1/1                0s 28ms/step
```

```
1/1                 0s 36ms/step
1/1                 0s 40ms/step
1/1                 0s 30ms/step
1/1                 0s 24ms/step
1/1                 0s 24ms/step
1/1                 0s 25ms/step
1/1                 0s 25ms/step
1/1                 0s 34ms/step
1/1                 0s 27ms/step
1/1                 0s 43ms/step
1/1                 0s 24ms/step
1/1                 0s 24ms/step
1/1                 0s 35ms/step
1/1                 0s 26ms/step
1/1                 0s 26ms/step
1/1                 0s 25ms/step
1/1                 0s 32ms/step
1/1                 0s 25ms/step
1/1                 0s 26ms/step
1/1                 0s 24ms/step
1/1                 0s 23ms/step
1/1                 0s 25ms/step
1/1                 0s 27ms/step
1/1                 0s 24ms/step
1/1                 0s 31ms/step
1/1                 0s 28ms/step
1/1                 0s 27ms/step
1/1                 0s 30ms/step
1/1                 0s 40ms/step
1/1                 0s 33ms/step
1/1                 0s 27ms/step
1/1                 0s 41ms/step
1/1                 0s 39ms/step
1/1                 0s 28ms/step
1/1                 0s 27ms/step
1/1                 0s 53ms/step
1/1                 0s 27ms/step
1/1                 0s 43ms/step
1/1                 0s 28ms/step
1/1                 0s 34ms/step
1/1                 0s 32ms/step
1/1                 0s 35ms/step
1/1                 0s 39ms/step
1/1                 0s 25ms/step
1/1                 0s 48ms/step
1/1                 0s 28ms/step
1/1                 0s 30ms/step
1/1                 0s 29ms/step
```

```
1/1              0s 43ms/step
1/1              0s 27ms/step
1/1              0s 36ms/step
1/1              0s 38ms/step
1/1              0s 41ms/step
1/1              0s 25ms/step
1/1              0s 26ms/step
1/1              0s 38ms/step
1/1              0s 33ms/step
1/1              0s 33ms/step
1/1              0s 44ms/step
1/1              0s 26ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 35ms/step
1/1              0s 27ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 26ms/step
1/1              0s 30ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 23ms/step
1/1              0s 24ms/step
1/1              0s 26ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 27ms/step
1/1              0s 25ms/step
1/1              0s 27ms/step
1/1              0s 27ms/step
1/1              0s 30ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 25ms/step
1/1              0s 25ms/step
1/1              0s 37ms/step
1/1              0s 25ms/step
1/1              0s 37ms/step
1/1              0s 27ms/step
1/1              0s 25ms/step
```

```
1/1              0s 27ms/step
1/1              0s 24ms/step
1/1              0s 25ms/step
1/1              0s 27ms/step
1/1              0s 25ms/step
1/1              0s 26ms/step
1/1              0s 29ms/step
1/1              0s 29ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 27ms/step
1/1              0s 53ms/step
1/1              0s 24ms/step
1/1              0s 31ms/step
1/1              0s 41ms/step
1/1              0s 27ms/step
1/1              0s 26ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 32ms/step
1/1              0s 28ms/step
1/1              0s 31ms/step
1/1              0s 29ms/step
1/1              0s 25ms/step
1/1              0s 24ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 24ms/step
1/1              0s 27ms/step
1/1              0s 26ms/step
1/1              0s 29ms/step
1/1              0s 26ms/step
1/1              0s 26ms/step
1/1              0s 28ms/step
1/1              0s 25ms/step
1/1              0s 35ms/step
1/1              0s 37ms/step
1/1              0s 39ms/step
1/1              0s 33ms/step
1/1              0s 57ms/step
1/1              0s 26ms/step
1/1              0s 53ms/step
1/1              0s 52ms/step
1/1              0s 52ms/step
1/1              0s 56ms/step
1/1              0s 33ms/step
1/1              0s 34ms/step
1/1              0s 28ms/step
```

```
1/1               0s 26ms/step
1/1               0s 42ms/step
1/1               0s 42ms/step
1/1               0s 27ms/step
1/1               0s 28ms/step
1/1               0s 30ms/step
1/1               0s 29ms/step
1/1               0s 44ms/step
1/1               0s 33ms/step
1/1               0s 26ms/step
1/1               0s 29ms/step
1/1               0s 32ms/step
1/1               0s 41ms/step
1/1               0s 42ms/step
1/1               0s 28ms/step
1/1               0s 27ms/step
1/1               0s 24ms/step
1/1               0s 25ms/step
1/1               0s 24ms/step
1/1               0s 24ms/step
1/1               0s 25ms/step
1/1               0s 35ms/step
1/1               0s 31ms/step
1/1               0s 26ms/step
1/1               0s 27ms/step
1/1               0s 24ms/step
1/1               0s 25ms/step
1/1               0s 27ms/step
1/1               0s 44ms/step
1/1               0s 34ms/step
1/1               0s 26ms/step
1/1               0s 25ms/step
1/1               0s 26ms/step
1/1               0s 25ms/step
1/1               0s 24ms/step
1/1               0s 39ms/step
1/1               0s 36ms/step
1/1               0s 37ms/step
1/1               0s 27ms/step
1/1               0s 27ms/step
1/1               0s 30ms/step
1/1               0s 24ms/step
1/1               0s 28ms/step
1/1               0s 25ms/step
1/1               0s 27ms/step
1/1               0s 33ms/step
1/1               0s 27ms/step
1/1               0s 25ms/step
```

```
1/1          0s 24ms/step
1/1          0s 24ms/step
1/1          0s 24ms/step
1/1          0s 34ms/step
1/1          0s 23ms/step
1/1          0s 24ms/step
1/1          0s 38ms/step
1/1          0s 24ms/step
1/1          0s 27ms/step
1/1          0s 25ms/step
1/1          0s 24ms/step
1/1          0s 36ms/step
1/1          0s 23ms/step
1/1          0s 29ms/step
1/1          0s 36ms/step
1/1          0s 24ms/step
1/1          0s 25ms/step
1/1          0s 24ms/step
1/1          0s 23ms/step
1/1          0s 23ms/step
1/1          0s 23ms/step
1/1          0s 25ms/step
1/1          0s 27ms/step
1/1          0s 27ms/step
1/1          0s 25ms/step
1/1          0s 24ms/step
1/1          0s 25ms/step
1/1          0s 27ms/step
1/1          0s 24ms/step
1/1          0s 25ms/step
1/1          0s 24ms/step
1/1          0s 25ms/step
1/1          0s 36ms/step
1/1          0s 38ms/step
1/1          0s 24ms/step
1/1          0s 34ms/step
1/1          0s 27ms/step
1/1          0s 25ms/step
1/1          0s 28ms/step
1/1          0s 26ms/step
1/1          0s 27ms/step
1/1          0s 24ms/step
1/1          0s 56ms/step
1/1          0s 45ms/step
1/1          0s 44ms/step
1/1          0s 34ms/step
1/1          0s 28ms/step
1/1          0s 49ms/step
```

```
1/1              0s 30ms/step
1/1              0s 40ms/step
1/1              0s 28ms/step
1/1              0s 39ms/step
1/1              0s 31ms/step
1/1              0s 54ms/step
1/1              0s 32ms/step
1/1              0s 27ms/step
1/1              0s 41ms/step
1/1              0s 42ms/step
1/1              0s 43ms/step
1/1              0s 35ms/step
1/1              0s 27ms/step
1/1              0s 35ms/step
1/1              0s 40ms/step
1/1              0s 26ms/step
1/1              0s 35ms/step
1/1              0s 43ms/step
1/1              0s 28ms/step
1/1              0s 38ms/step
1/1              0s 35ms/step
1/1              0s 26ms/step
1/1              0s 24ms/step
1/1              0s 26ms/step
1/1              0s 29ms/step
Average Energy Consumption with RL: 93.65 Wh
Average Energy Savings per Sample: 5.69 Wh
```

By addressing the scaling issues, correctly handling energy consumption, and ensuring that actions influence energy usage through a trained energy prediction model, your Actor-Critic reinforcement learning algorithm should effectively learn to adjust temperature settings to optimize energy consumption while maintaining comfort.

Key Takeaways: Separate Scaling: Always ensure that features directly involved in reward calculations are not improperly scaled, as this can distort the reward signal and hinder learning.

Environment Simulation: Actions must have a meaningful impact on the environment. By using a trained energy prediction model, you allow the RL agent's actions to influence energy consumption, enabling effective learning.

Reward Function Design: Carefully design the reward function to balance multiple objectives (e.g., comfort and energy efficiency) by appropriately weighting different components.

Model Training and Evaluation: Continuously monitor and evaluate your models to ensure convergence and effectiveness. Use visualizations to understand learning dynamics and policy behavior.

Iterative Refinement: RL algorithms often require iterative tuning of hyperparameters, reward functions, and environment simulations to achieve optimal performance.

[ ]: