

Implement the DQN and DDQN algorithm

0.0.1 Group ID: DRL Group 150

0.0.2 Group Members Name with Student ID:

1. HEMANT KUMAR PARAKH (2023aa05741)
2. SUSHIL KUMAR (2023aa05849)
3. NAGINENI SATISH BABU (2023aa05585)
4. JITENDRA KUMAR (2023aa05198)

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_hg: 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 = $\text{abs}(T_i - T_{\text{target}})$

Deviations=[$\text{abs}(22-22)$, $\text{abs}(21-22)$, $\text{abs}(20-22)$, $\text{abs}(22-22)$, $\text{abs}(21-22)$, $\text{abs}(20-22)$, $\text{abs}(23-22)$, $\text{abs}(21-22)$, $\text{abs}(22-22)$]

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 : Implement the DQN and DDQN algorithm

1.0.1 Imports

```
[4]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
import gc
import random

# PyTorch imports
import torch
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

1.0.2 Installation of Pytorch libs

```
[5]: !pip install torchinfo
from torchinfo import summary
```

Requirement already satisfied: torchinfo in /usr/local/lib/python3.10/dist-packages (1.8.0)

1.0.3 Mount google drive to load dataset

```
[6]: from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

1. Load Dataset
2. Check and fill missing values with mean
3. Defining X and y Data as features and target respectively

4. Data normalization and scalling
5. Splitting data for train and test

```
[7]: ##### Load the dataset
try:
    # Load the dataset
    file_path = '/content/energydata_complete.csv'
    df = pd.read_csv(file_path)
except:
    # Load the dataset
    file_path = '/content/drive/MyDrive/Colab Notebooks/DRL-assignment-2/
    ↪energydata_complete.csv'
    df = pd.read_csv(file_path)

print("dataset size = ", df.shape)

#data = df.sample(frac=0.01, random_state=42) # Sampling for quicker processing
#print("Reduced % of dataset size = ", data.shape)

data = df.copy()

# Check and replace missing values
numeric_data = data.select_dtypes(include=np.number)
data.fillna(numeric_data.mean(), inplace=True)

# Preprocess the data
features = ['T1', 'T2', 'T3', 'T4', 'T5', 'T6', 'T7', 'T8', 'T9',
            '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']

X = data[features]
y = data[target]

# Normalize them with Standard Scaler
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Split the data to training and testing sets (80% for training, 20% for
↪testing)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
↪random_state=0)
print(f"X_train shape: {X_train.shape}, y_train shape: {y_train.shape}")
print(f"X_test shape: {X_test.shape}, y_test shape: {y_test.shape}")
```

dataset size = (19735, 29)

X_train shape: (15788, 22), y_train shape: (15788, 1)
X_test shape: (3947, 22), y_test shape: (3947, 1)

#Data Visulization

```
[8]: import matplotlib.pyplot as plt

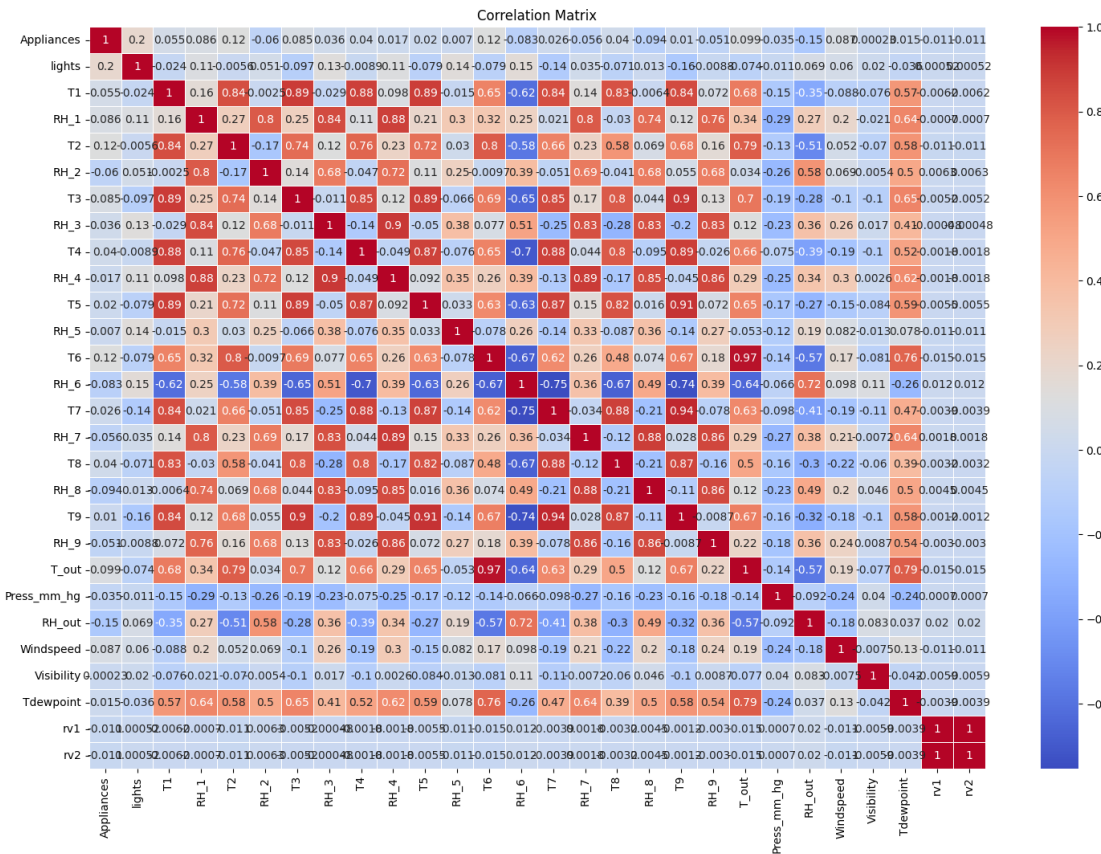
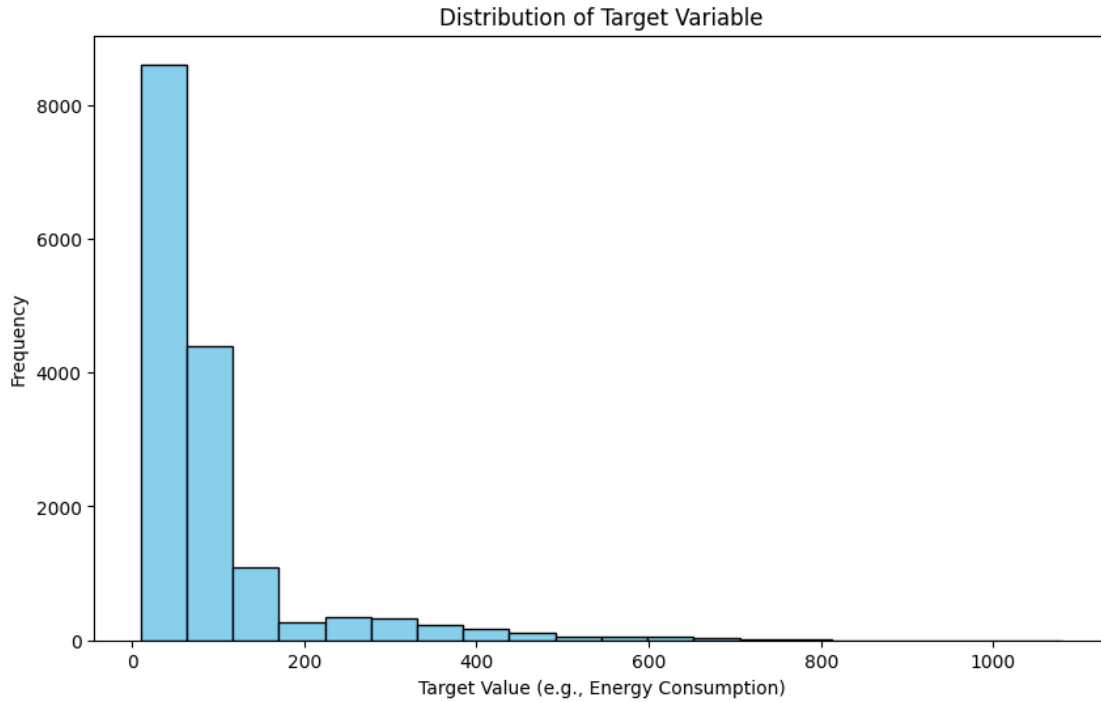
def plot_target_distribution(y):
    plt.figure(figsize=(10, 6))
    plt.hist(y, bins=20, color='skyblue', edgecolor='black')
    plt.title('Distribution of Target Variable')
    plt.xlabel('Target Value (e.g., Energy Consumption)')
    plt.ylabel('Frequency')
    plt.show()

# Assuming y_train contains the target values for the training set
plot_target_distribution(y_train)

import seaborn as sns

def plot_correlation_matrix(data):
    plt.figure(figsize=(18, 12))
    correlation_matrix = data.drop('date', axis=1).corr()
    sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
    plt.title('Correlation Matrix')
    plt.show()

# Assuming 'data' is the complete dataset including the target column
plot_correlation_matrix(data)
```



1.0.4 Create an EnergyConsumption Environment (1 M)

Creating Energy Consumption environment with required action and rewards function

```
[9]: ### write your code below this line

max_steps_per_episode = 200 # or any reasonable number

# Normalization of reward with mean and SD
class NormalizedRewards:
    def __init__(self, alpha=0.001, beta=0.001):
        self.rewards = []
        self.alpha = alpha # Smoothing factor for mean
        self.beta = beta # Smoothing factor for std
        self.mean_reward = 0
        self.std_reward = 1

    def update(self, reward):
        # Maintain a running estimate of the mean and std deviation
        self.mean_reward = self.alpha * reward + (1 - self.alpha) * self.
↪mean_reward
        self.std_reward = self.beta * np.abs(reward - self.mean_reward) + (1 -
↪self.beta) * self.std_reward

    def normalize(self, reward):
        self.update(reward)
        # Normalize the reward using the running mean and std
        return (reward - self.mean_reward) / (self.std_reward + 1e-5)

# Energy Consumptio Environment
class EnergyConsumptionEnv:
    def __init__(self, data, target_temp=22, max_steps=200):
        # Initializing the environment with data
        self.data = data
        self.target_temp = target_temp
        self.max_steps = max_steps # Set the max steps per episode
        # Define the state space (features) and action space
        self.state_space = data.shape[1] # State space size (number of
↪features)
        self.action_space = 3 # Three possible actions (decrease, maintain,
↪increase temp)
        self.reward_normalizer = NormalizedRewards(alpha=0.001, beta=0.001)

        # Keep track of the current step in the dataset
        self.current_step = 0
```

```

# reset environment
def reset(self):
    """Resets the environment to the initial state"""
    self.current_step = 0
    return self.data.iloc[self.current_step, :].values

# Step function to define the actions and move to next state
def step(self, action):
    """Take an action (0, 1, 2) and move to the next state"""

    # Extract the current state from the dataset
    #print(f"Current step: {self.current_step}, Data length: {len(self.
↪data)}")
    if 0 <= self.current_step < len(self.data):
        current_state = self.data.iloc[self.current_step, :].values
    else:
        #print(f"Current step after: {self.current_step}, Data length:
↪{len(self.data)}")
        current_state = None
        return None, 0, True

    # Determine the next step (either increment by 1 or stop at the end of
↪the data)
    next_step = min(self.current_step + 1, len(self.data) - 1)
    next_state = self.data.iloc[next_step, :].values

    # Apply the action to the first 9 temperature features (T1-T9)
    for i in range(9):
        if action == 0: # Decrease temperature
            next_state[i] -= 1
        elif action == 2: # Increase temperature
            next_state[i] += 1

    # Calculate the reward based on energy savings and temperature deviation
    reward = self.calculate_reward(current_state, next_state)

    # Increment the step count after taking an action
    if self.current_step < len(self.data):
        self.current_step += 1
    else:
        # Handle end of episode (e.g., set done = True)
        done = True

    # Determine if the episode is done (either at the end of data or
↪maximum steps reached)

```



```

        done = self.current_step >= len(self.data) - 1 or self.current_step >= self.max_steps

        # Debugging output to track the termination conditions
        #print(f"Step {self.current_step}, Max steps: {self.max_steps}, Data length: {len(self.data)}")
        #print(f"Step taken, Done: {done}, Step count: {self.current_step}/{self.max_steps}")

        # Return the next state, the reward, and the done flag
        return next_state, reward, done

    # Reward calculations
    def calculate_reward(self, current_state, next_state):
        """Calculates the reward based on energy savings and comfort"""
        if current_state is None or next_state is None:
            return 0 # Handle invalid state with a default reward

        # Proceed with normal reward calculation if state is valid
        # Extract energy consumption (last column of the state)
        energy_before = current_state[-1]
        energy_after = next_state[-1]
        energy_savings = current_state[-1] - next_state[-1]

        # Calculate temperature deviations from the target
        deviations = np.abs(next_state[:9] - self.target_temp) # Assume T1-T9 are first 9 columns
        comfort_penalty = np.sum(deviations)

        # Reward is a combination of energy savings and comfort penalty
        reward = energy_savings - comfort_penalty

        normalized_reward = self.reward_normalizer.normalize(reward)
        return normalized_reward

    def action_space_sample(self):
        """Randomly sample an action (0: decrease, 1: maintain, 2: increase)"""
        return random.choice([0, 1, 2])

    # state space size
    def get_state_space_size(self):
        return self.state_space

    # Action space size
    def get_action_space_size(self):
        return self.action_space

```

```
# Initialize environment
env = EnergyConsumptionEnv(data=X_train, max_steps=max_steps_per_episode)
```

1.0.5 Print state space and action space (0.5 M)

```
[10]: ### write your code below this line
# Print the state and action space sizes
print("State space size:", env.get_state_space_size())
print("Action space size:", env.get_action_space_size())
```

State space size: 22

Action space size: 3

1.0.6 Clearly define the parameters used for training an AI agent. (1 M)

- Number of episodes
- Max capacity of replay memory
- Batch size
- Period of Q target network updates
- Discount factor for future rewards
- Initial value for epsilon of the e-greedy
- Final value for epsilon of the e-greedy
- Learning rate of ADAM optimizer, and etc.

```
[11]: ### write your code below this line
# Define training parameters
num_episodes = 500
max_replay_memory = 10000
batch_size = 64
target_update_period = 100 # Update target network every 100 steps
gamma = 0.99 # Discount factor
epsilon_start = 1.0 # Starting value of epsilon for epsilon-greedy policy
epsilon_end = 0.1 # Final value of epsilon
epsilon_decay = 0.995 # Epsilon decay rate
learning_rate = 0.001 # Learning rate for ADAM optimizer
lr_decay = 0.99 # Decay learning rate by 1% every adjustment
td_error = 1.0
```

Define the separate functions for DecreaseTemperature, IncreaseTemperature and MaintainCurrentTemperature actions. (1.5 M)

```
[12]: ### write your code below this line
def decrease_temperature(state):
    for i in range(9): # Assuming T1 to T9 are the first 9 columns
        state[i] = max(state[i] - 1, -10) # Clamp to -10°C minimum
    return state
```

```

def increase_temperature(state):
    for i in range(9):
        state[i] = min(state[i] + 1, 30) # Clamp to 30°C maximum
    return state

def maintain_temperature(state):
    # No changes to temperature
    return state

```

Implement a replay buffer for storing the experiences. (0.5 M) Defining the Prioritized Replay Buffer to improve the learning efficiency by assigns a priority score to each experience

```

[13]: ### write your code below this line
from collections import deque

class PrioritizedReplayBuffer:
    def __init__(self, max_size, alpha=0.6):
        self.buffer = deque(maxlen=max_size)
        self.priorities = deque(maxlen=max_size)
        self.alpha = alpha

    def add(self, experience, td_error):
        self.buffer.append(experience)
        max_priority = max(self.priorities) if self.priorities else 1.0
        self.priorities.append(max_priority)

    def sample(self, batch_size, beta=0.4):
        priorities = np.array(self.priorities, dtype=np.float32)
        probs = priorities ** self.alpha
        probs /= probs.sum()

        indices = np.random.choice(len(self.buffer), batch_size, p=probs)
        samples = [self.buffer[i] for i in indices]

        total = len(self.buffer)
        weights = (total * probs[indices]) ** (-beta)
        weights /= weights.max()

        return samples, indices, weights

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

    def update_priorities(self, indices, td_errors):
        for i, td_error in zip(indices, td_errors):
            self.priorities[i] = abs(td_error) + 1e-5 # Small value to prevent
            ↪ 0 priority

```

```
# Initialize replay buffer
replay_buffer = PrioritizedReplayBuffer(max_size=max_replay_memory)
```

1.0.7 Design network DQN (0.5 M)

Designing the DQN model with fully connected layers

```
[14]: ### write your code below this line
class DQN(nn.Module):
    def __init__(self, state_size, action_size):
        super(DQN, self).__init__()
        self.fc1 = nn.Linear(state_size, 64)
        self.fc2 = nn.Linear(64, 128)
        self.fc3 = nn.Linear(128, action_size)

    def forward(self, x):
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        return self.fc3(x)

# Initialize DQN model and optimizer
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
dqn_model = DQN(env.get_state_space_size(), env.get_action_space_size()).
    ↪to(device)
target_dqn_model = DQN(env.get_state_space_size(), env.get_action_space_size()).
    ↪to(device)
target_dqn_model.load_state_dict(dqn_model.state_dict())
target_dqn_model.eval() # Target network in eval mode
optimizer = optim.Adam(dqn_model.parameters(), lr=learning_rate)
loss_fn = nn.MSELoss()

# print DQN model summary
print(dqn_model)
summary(target_dqn_model, input_size=(batch_size, env.get_state_space_size()))
```

```
DQN(
  (fc1): Linear(in_features=22, out_features=64, bias=True)
  (fc2): Linear(in_features=64, out_features=128, bias=True)
  (fc3): Linear(in_features=128, out_features=3, bias=True)
)
```

```
[14]: =====
=====
Layer (type:depth-idx)                Output Shape                Param #
=====
```

```

=====
DQN                                [64, 3]                                --
  Linear: 1-1                      [64, 64]                                1,472
  Linear: 1-2                      [64, 128]                               8,320
  Linear: 1-3                      [64, 3]                                387
=====
=====
Total params: 10,179
Trainable params: 10,179
Non-trainable params: 0
Total mult-adds (M): 0.65
=====
=====
Input size (MB): 0.01
Forward/backward pass size (MB): 0.10
Params size (MB): 0.04
Estimated Total Size (MB): 0.15
=====
=====

```

1.0.8 Training Iterations (0.5 M)

Note: print all the episodes with the values of investment and buying. (if not printed then -1 will be done.)

```

[15]: ### write your code below this line
      # Training loop for DQN in PyTorch
      dqn_rewards = []
      epsilon = epsilon_start
      target_update_freq = 10

      # Initialize counters for actions for both DQN and DDQN
      dqn_action_counts = {'decrease': 0, 'maintain': 0, 'increase': 0}

      # Add learning rate decay to the optimizer
      lr_scheduler = torch.optim.lr_scheduler.StepLR(optimizer, step_size=100,
      ↪gamma=lr_decay)

      # Iterate through number of episodes defined
      for episode in range(num_episodes):
          state = env.reset()
          state = torch.FloatTensor(state).to(device).unsqueeze(0) # Convert state
          ↪to tensor
          total_reward = 0
          total_investment = 0 # Track total investment (actions)
          total_buying = 0 # Track total buying (rewards)
          done = False

```

```

step_count = 0

while not done:
    if np.random.rand() < epsilon:
        action = env.action_space_sample()
    else:
        with torch.no_grad():
            q_values = dqn_model(state)
            action = torch.argmax(q_values).item()

    # Track action counts for DQN
    if action == 0:
        dqn_action_counts['decrease'] += 1
    elif action == 1:
        dqn_action_counts['maintain'] += 1
    else:
        dqn_action_counts['increase'] += 1

    next_state, reward, done = env.step(action)
    next_state = torch.FloatTensor(next_state).to(device).unsqueeze(0)
    reward = torch.tensor([reward], dtype=torch.float32, device=device)

    # Track investment and buying
    total_investment += action # Assuming action represents some form of ↵
    ↵energy adjustment
    total_buying += reward.item() # Reward represents the "buying" or ↵
    ↵benefit

    # Store experience in replay buffer
    replay_buffer.add((state, action, reward, next_state, done), td_error)

    # Sample from replay buffer and train
    if replay_buffer.size() >= batch_size:
        experiences, indices, weights = replay_buffer.sample(batch_size)
        states, actions, rewards, next_states, dones = zip(*experiences)
        states = torch.cat(states).to(device)
        next_states = torch.cat(next_states).to(device)
        actions = torch.tensor(actions, device=device).unsqueeze(1)
        rewards = torch.cat(rewards).to(device)

        q_values = dqn_model(states).gather(1, actions)
        next_q_values = target_dqn_model(next_states).max(1)[0].detach()
        targets = rewards + (gamma * next_q_values * (1 - torch.
    ↵tensor(dones, dtype=torch.float32, device=device)))

        loss = loss_fn(q_values.squeeze(), targets)

```

```

optimizer.zero_grad()
loss.backward()
optimizer.step()

# Calculate TD error and update priorities
td_errors = (targets - q_values.squeeze()).detach().cpu().numpy()
replay_buffer.update_priorities(indices, td_errors)

lr_scheduler.step() # Apply learning rate decay

state = next_state
total_reward += reward.item()
step_count += 1

dqn_rewards.append(total_reward)

# Decay epsilon
epsilon = max(epsilon_end, epsilon * epsilon_decay)

# Print investment and buying at the end of each episode
print(f"Episode {episode + 1}/{num_episodes}, Investment (total actions):␣
↪{total_investment}, Buying (total rewards): {total_buying}, Total Reward:␣
↪{total_reward}")

# Update the target network every 'target_update_freq' episodes
if episode % target_update_freq == 0:
    target_dqn_model.load_state_dict(dqn_model.state_dict()) # Sync the␣
↪target network weights
    print(f"Target network updated at episode {episode}")

print("DQN Training completed!")

```

```

Episode 1/500, Investment (total actions): 214, Buying (total rewards):
-1981.2049784064293, Total Reward: -1981.2049784064293
Target network updated at episode 0
Episode 2/500, Investment (total actions): 217, Buying (total rewards):
-636.2258600890636, Total Reward: -636.2258600890636
Episode 3/500, Investment (total actions): 188, Buying (total rewards):
-401.2356416583061, Total Reward: -401.2356416583061
Episode 4/500, Investment (total actions): 215, Buying (total rewards):
-283.2553413398564, Total Reward: -283.2553413398564
Episode 5/500, Investment (total actions): 205, Buying (total rewards):
-230.5664141587913, Total Reward: -230.5664141587913
Episode 6/500, Investment (total actions): 192, Buying (total rewards):
-193.91695779934525, Total Reward: -193.91695779934525
Episode 7/500, Investment (total actions): 205, Buying (total rewards):
-164.63762766873697, Total Reward: -164.63762766873697

```

Episode 8/500, Investment (total actions): 226, Buying (total rewards):
 -127.31066384265432, Total Reward: -127.31066384265432
 Episode 9/500, Investment (total actions): 205, Buying (total rewards):
 -109.07764837192371, Total Reward: -109.07764837192371
 Episode 10/500, Investment (total actions): 199, Buying (total rewards):
 -99.27112793829292, Total Reward: -99.27112793829292
 Episode 11/500, Investment (total actions): 195, Buying (total rewards):
 -89.49048534687608, Total Reward: -89.49048534687608
 Target network updated at episode 10
 Episode 12/500, Investment (total actions): 215, Buying (total rewards):
 -69.8412708779797, Total Reward: -69.8412708779797
 Episode 13/500, Investment (total actions): 214, Buying (total rewards):
 -70.9855862716795, Total Reward: -70.9855862716795
 Episode 14/500, Investment (total actions): 211, Buying (total rewards):
 -64.55592238157988, Total Reward: -64.55592238157988
 Episode 15/500, Investment (total actions): 218, Buying (total rewards):
 -50.551689642015845, Total Reward: -50.551689642015845
 Episode 16/500, Investment (total actions): 206, Buying (total rewards):
 -48.93908666144125, Total Reward: -48.93908666144125
 Episode 17/500, Investment (total actions): 187, Buying (total rewards):
 -48.27865776454564, Total Reward: -48.27865776454564
 Episode 18/500, Investment (total actions): 210, Buying (total rewards):
 -44.37012484204024, Total Reward: -44.37012484204024
 Episode 19/500, Investment (total actions): 209, Buying (total rewards):
 -38.00706938840449, Total Reward: -38.00706938840449
 Episode 20/500, Investment (total actions): 219, Buying (total rewards):
 -41.38527253828943, Total Reward: -41.38527253828943
 Episode 21/500, Investment (total actions): 209, Buying (total rewards):
 -38.73722045496106, Total Reward: -38.73722045496106
 Target network updated at episode 20
 Episode 22/500, Investment (total actions): 197, Buying (total rewards):
 -40.970228425459936, Total Reward: -40.970228425459936
 Episode 23/500, Investment (total actions): 191, Buying (total rewards):
 -37.21751586534083, Total Reward: -37.21751586534083
 Episode 24/500, Investment (total actions): 203, Buying (total rewards):
 -37.378217115066946, Total Reward: -37.378217115066946
 Episode 25/500, Investment (total actions): 206, Buying (total rewards):
 -32.93518843315542, Total Reward: -32.93518843315542
 Episode 26/500, Investment (total actions): 194, Buying (total rewards):
 -31.218214023392648, Total Reward: -31.218214023392648
 Episode 27/500, Investment (total actions): 210, Buying (total rewards):
 -30.924216041108593, Total Reward: -30.924216041108593
 Episode 28/500, Investment (total actions): 200, Buying (total rewards):
 -37.09972753468901, Total Reward: -37.09972753468901
 Episode 29/500, Investment (total actions): 201, Buying (total rewards):
 -32.41859648388345, Total Reward: -32.41859648388345
 Episode 30/500, Investment (total actions): 206, Buying (total rewards):
 -38.537892242718954, Total Reward: -38.537892242718954

Episode 31/500, Investment (total actions): 184, Buying (total rewards):
 -38.09076880663633, Total Reward: -38.09076880663633
 Target network updated at episode 30
 Episode 32/500, Investment (total actions): 209, Buying (total rewards):
 -36.652432948350906, Total Reward: -36.652432948350906
 Episode 33/500, Investment (total actions): 184, Buying (total rewards):
 -34.81449181586504, Total Reward: -34.81449181586504
 Episode 34/500, Investment (total actions): 183, Buying (total rewards):
 -32.277464495040476, Total Reward: -32.277464495040476
 Episode 35/500, Investment (total actions): 208, Buying (total rewards):
 -25.223464336246252, Total Reward: -25.223464336246252
 Episode 36/500, Investment (total actions): 203, Buying (total rewards):
 -24.870014983578585, Total Reward: -24.870014983578585
 Episode 37/500, Investment (total actions): 178, Buying (total rewards):
 -22.624588179402053, Total Reward: -22.624588179402053
 Episode 38/500, Investment (total actions): 216, Buying (total rewards):
 -11.693711408413947, Total Reward: -11.693711408413947
 Episode 39/500, Investment (total actions): 212, Buying (total rewards):
 -13.497208265587687, Total Reward: -13.497208265587687
 Episode 40/500, Investment (total actions): 193, Buying (total rewards):
 -14.493935420177877, Total Reward: -14.493935420177877
 Episode 41/500, Investment (total actions): 206, Buying (total rewards):
 -22.473642718046904, Total Reward: -22.473642718046904
 Target network updated at episode 40
 Episode 42/500, Investment (total actions): 221, Buying (total rewards):
 -22.214567615184933, Total Reward: -22.214567615184933
 Episode 43/500, Investment (total actions): 213, Buying (total rewards):
 -21.40008439309895, Total Reward: -21.40008439309895
 Episode 44/500, Investment (total actions): 200, Buying (total rewards):
 -24.822568918112665, Total Reward: -24.822568918112665
 Episode 45/500, Investment (total actions): 209, Buying (total rewards):
 -22.31872812844813, Total Reward: -22.31872812844813
 Episode 46/500, Investment (total actions): 193, Buying (total rewards):
 -16.24942352529615, Total Reward: -16.24942352529615
 Episode 47/500, Investment (total actions): 207, Buying (total rewards):
 -16.70508108008653, Total Reward: -16.70508108008653
 Episode 48/500, Investment (total actions): 203, Buying (total rewards):
 -16.023190993466415, Total Reward: -16.023190993466415
 Episode 49/500, Investment (total actions): 197, Buying (total rewards):
 -14.117345662787557, Total Reward: -14.117345662787557
 Episode 50/500, Investment (total actions): 225, Buying (total rewards):
 -16.03371547907591, Total Reward: -16.03371547907591
 Episode 51/500, Investment (total actions): 200, Buying (total rewards):
 -11.122589704580605, Total Reward: -11.122589704580605
 Target network updated at episode 50
 Episode 52/500, Investment (total actions): 223, Buying (total rewards):
 -11.152924083173275, Total Reward: -11.152924083173275
 Episode 53/500, Investment (total actions): 209, Buying (total rewards):

-13.580774372443557, Total Reward: -13.580774372443557
 Episode 54/500, Investment (total actions): 213, Buying (total rewards):
 -9.230048330500722, Total Reward: -9.230048330500722
 Episode 55/500, Investment (total actions): 216, Buying (total rewards):
 -7.623983905650675, Total Reward: -7.623983905650675
 Episode 56/500, Investment (total actions): 194, Buying (total rewards):
 -16.166423376649618, Total Reward: -16.166423376649618
 Episode 57/500, Investment (total actions): 197, Buying (total rewards):
 -19.824616035446525, Total Reward: -19.824616035446525
 Episode 58/500, Investment (total actions): 201, Buying (total rewards):
 -22.44044304289855, Total Reward: -22.44044304289855
 Episode 59/500, Investment (total actions): 206, Buying (total rewards):
 -17.688867276534438, Total Reward: -17.688867276534438
 Episode 60/500, Investment (total actions): 215, Buying (total rewards):
 -17.343879517167807, Total Reward: -17.343879517167807
 Episode 61/500, Investment (total actions): 221, Buying (total rewards):
 -18.02703483426012, Total Reward: -18.02703483426012
 Target network updated at episode 60
 Episode 62/500, Investment (total actions): 199, Buying (total rewards):
 -23.64281091396697, Total Reward: -23.64281091396697
 Episode 63/500, Investment (total actions): 197, Buying (total rewards):
 -20.909161592833698, Total Reward: -20.909161592833698
 Episode 64/500, Investment (total actions): 210, Buying (total rewards):
 -19.122626438736916, Total Reward: -19.122626438736916
 Episode 65/500, Investment (total actions): 212, Buying (total rewards):
 -12.842979166656733, Total Reward: -12.842979166656733
 Episode 66/500, Investment (total actions): 193, Buying (total rewards):
 -10.791303811594844, Total Reward: -10.791303811594844
 Episode 67/500, Investment (total actions): 206, Buying (total rewards):
 -11.638829258736223, Total Reward: -11.638829258736223
 Episode 68/500, Investment (total actions): 205, Buying (total rewards):
 -5.041317927534692, Total Reward: -5.041317927534692
 Episode 69/500, Investment (total actions): 208, Buying (total rewards):
 -3.925825196551159, Total Reward: -3.925825196551159
 Episode 70/500, Investment (total actions): 199, Buying (total rewards):
 -4.905112245120108, Total Reward: -4.905112245120108
 Episode 71/500, Investment (total actions): 181, Buying (total rewards):
 -2.096798966638744, Total Reward: -2.096798966638744
 Target network updated at episode 70
 Episode 72/500, Investment (total actions): 196, Buying (total rewards):
 -1.7345474250614643, Total Reward: -1.7345474250614643
 Episode 73/500, Investment (total actions): 208, Buying (total rewards):
 -7.339372405316681, Total Reward: -7.339372405316681
 Episode 74/500, Investment (total actions): 183, Buying (total rewards):
 -10.231567774899304, Total Reward: -10.231567774899304
 Episode 75/500, Investment (total actions): 223, Buying (total rewards):
 -5.038125501712784, Total Reward: -5.038125501712784
 Episode 76/500, Investment (total actions): 220, Buying (total rewards):

-7.547704486889415, Total Reward: -7.547704486889415
 Episode 77/500, Investment (total actions): 219, Buying (total rewards):
 -8.920074798166752, Total Reward: -8.920074798166752
 Episode 78/500, Investment (total actions): 216, Buying (total rewards):
 -12.83939142525196, Total Reward: -12.83939142525196
 Episode 79/500, Investment (total actions): 205, Buying (total rewards):
 -11.91263661114499, Total Reward: -11.91263661114499
 Episode 80/500, Investment (total actions): 201, Buying (total rewards):
 -6.684296603547409, Total Reward: -6.684296603547409
 Episode 81/500, Investment (total actions): 213, Buying (total rewards):
 -10.082698370055368, Total Reward: -10.082698370055368
 Target network updated at episode 80
 Episode 82/500, Investment (total actions): 188, Buying (total rewards):
 -11.529255233937874, Total Reward: -11.529255233937874
 Episode 83/500, Investment (total actions): 194, Buying (total rewards):
 -11.053578786551952, Total Reward: -11.053578786551952
 Episode 84/500, Investment (total actions): 200, Buying (total rewards):
 -10.368332887301221, Total Reward: -10.368332887301221
 Episode 85/500, Investment (total actions): 201, Buying (total rewards):
 -7.964766537770629, Total Reward: -7.964766537770629
 Episode 86/500, Investment (total actions): 204, Buying (total rewards):
 -8.222975172102451, Total Reward: -8.222975172102451
 Episode 87/500, Investment (total actions): 201, Buying (total rewards):
 -5.79448468843475, Total Reward: -5.79448468843475
 Episode 88/500, Investment (total actions): 226, Buying (total rewards):
 -5.961195134674199, Total Reward: -5.961195134674199
 Episode 89/500, Investment (total actions): 205, Buying (total rewards):
 -13.3767089988105, Total Reward: -13.3767089988105
 Episode 90/500, Investment (total actions): 225, Buying (total rewards):
 -15.699158186092973, Total Reward: -15.699158186092973
 Episode 91/500, Investment (total actions): 226, Buying (total rewards):
 -12.497397654224187, Total Reward: -12.497397654224187
 Target network updated at episode 90
 Episode 92/500, Investment (total actions): 205, Buying (total rewards):
 -9.761857035569847, Total Reward: -9.761857035569847
 Episode 93/500, Investment (total actions): 206, Buying (total rewards):
 -9.539670481346548, Total Reward: -9.539670481346548
 Episode 94/500, Investment (total actions): 201, Buying (total rewards):
 -12.182046078145504, Total Reward: -12.182046078145504
 Episode 95/500, Investment (total actions): 211, Buying (total rewards):
 -9.578641256317496, Total Reward: -9.578641256317496
 Episode 96/500, Investment (total actions): 177, Buying (total rewards):
 -9.660644083283842, Total Reward: -9.660644083283842
 Episode 97/500, Investment (total actions): 209, Buying (total rewards):
 -8.455555256921798, Total Reward: -8.455555256921798
 Episode 98/500, Investment (total actions): 197, Buying (total rewards):
 -5.804830510169268, Total Reward: -5.804830510169268
 Episode 99/500, Investment (total actions): 198, Buying (total rewards):

-1.765309865353629, Total Reward: -1.765309865353629
 Episode 100/500, Investment (total actions): 203, Buying (total rewards):
 -1.8058401476591825, Total Reward: -1.8058401476591825
 Episode 101/500, Investment (total actions): 190, Buying (total rewards):
 -8.526296810014173, Total Reward: -8.526296810014173
 Target network updated at episode 100
 Episode 102/500, Investment (total actions): 175, Buying (total rewards):
 -10.93909527733922, Total Reward: -10.93909527733922
 Episode 103/500, Investment (total actions): 216, Buying (total rewards):
 -12.38056241767481, Total Reward: -12.38056241767481
 Episode 104/500, Investment (total actions): 218, Buying (total rewards):
 -11.21940628066659, Total Reward: -11.21940628066659
 Episode 105/500, Investment (total actions): 203, Buying (total rewards):
 -9.723720418289304, Total Reward: -9.723720418289304
 Episode 106/500, Investment (total actions): 208, Buying (total rewards):
 -13.85469943494536, Total Reward: -13.85469943494536
 Episode 107/500, Investment (total actions): 214, Buying (total rewards):
 -15.967449620366096, Total Reward: -15.967449620366096
 Episode 108/500, Investment (total actions): 206, Buying (total rewards):
 -19.963932475540787, Total Reward: -19.963932475540787
 Episode 109/500, Investment (total actions): 184, Buying (total rewards):
 -19.638077267678455, Total Reward: -19.638077267678455
 Episode 110/500, Investment (total actions): 211, Buying (total rewards):
 -18.52794071077369, Total Reward: -18.52794071077369
 Episode 111/500, Investment (total actions): 200, Buying (total rewards):
 -21.472274221479893, Total Reward: -21.472274221479893
 Target network updated at episode 110
 Episode 112/500, Investment (total actions): 167, Buying (total rewards):
 -23.13771045161411, Total Reward: -23.13771045161411
 Episode 113/500, Investment (total actions): 187, Buying (total rewards):
 -25.303789602126926, Total Reward: -25.303789602126926
 Episode 114/500, Investment (total actions): 226, Buying (total rewards):
 -30.062851656693965, Total Reward: -30.062851656693965
 Episode 115/500, Investment (total actions): 222, Buying (total rewards):
 -28.704550015507266, Total Reward: -28.704550015507266
 Episode 116/500, Investment (total actions): 195, Buying (total rewards):
 -30.52608173713088, Total Reward: -30.52608173713088
 Episode 117/500, Investment (total actions): 210, Buying (total rewards):
 -25.015719322604127, Total Reward: -25.015719322604127
 Episode 118/500, Investment (total actions): 198, Buying (total rewards):
 -26.977206964045763, Total Reward: -26.977206964045763
 Episode 119/500, Investment (total actions): 225, Buying (total rewards):
 -26.37998253433034, Total Reward: -26.37998253433034
 Episode 120/500, Investment (total actions): 225, Buying (total rewards):
 -25.260652666911483, Total Reward: -25.260652666911483
 Episode 121/500, Investment (total actions): 205, Buying (total rewards):
 -22.78244245913811, Total Reward: -22.78244245913811
 Target network updated at episode 120

Episode 122/500, Investment (total actions): 209, Buying (total rewards):
 -21.12289122887887, Total Reward: -21.12289122887887
 Episode 123/500, Investment (total actions): 193, Buying (total rewards):
 -20.610943216830492, Total Reward: -20.610943216830492
 Episode 124/500, Investment (total actions): 188, Buying (total rewards):
 -19.660683439142304, Total Reward: -19.660683439142304
 Episode 125/500, Investment (total actions): 214, Buying (total rewards):
 -20.273151034489274, Total Reward: -20.273151034489274
 Episode 126/500, Investment (total actions): 187, Buying (total rewards):
 -19.15360910841264, Total Reward: -19.15360910841264
 Episode 127/500, Investment (total actions): 193, Buying (total rewards):
 -17.987045705784112, Total Reward: -17.987045705784112
 Episode 128/500, Investment (total actions): 192, Buying (total rewards):
 -19.93202494457364, Total Reward: -19.93202494457364
 Episode 129/500, Investment (total actions): 191, Buying (total rewards):
 -16.86332238931209, Total Reward: -16.86332238931209
 Episode 130/500, Investment (total actions): 193, Buying (total rewards):
 -16.210006826207973, Total Reward: -16.210006826207973
 Episode 131/500, Investment (total actions): 192, Buying (total rewards):
 -14.547193909063935, Total Reward: -14.547193909063935
 Target network updated at episode 130
 Episode 132/500, Investment (total actions): 185, Buying (total rewards):
 -15.662427350878716, Total Reward: -15.662427350878716
 Episode 133/500, Investment (total actions): 171, Buying (total rewards):
 -17.176721275784075, Total Reward: -17.176721275784075
 Episode 134/500, Investment (total actions): 173, Buying (total rewards):
 -18.445102468598634, Total Reward: -18.445102468598634
 Episode 135/500, Investment (total actions): 190, Buying (total rewards):
 -16.43648361042142, Total Reward: -16.43648361042142
 Episode 136/500, Investment (total actions): 197, Buying (total rewards):
 -16.359932395746, Total Reward: -16.359932395746
 Episode 137/500, Investment (total actions): 213, Buying (total rewards):
 -17.12315607018536, Total Reward: -17.12315607018536
 Episode 138/500, Investment (total actions): 175, Buying (total rewards):
 -18.224601475521922, Total Reward: -18.224601475521922
 Episode 139/500, Investment (total actions): 205, Buying (total rewards):
 -16.677486626431346, Total Reward: -16.677486626431346
 Episode 140/500, Investment (total actions): 192, Buying (total rewards):
 -18.075752406846732, Total Reward: -18.075752406846732
 Episode 141/500, Investment (total actions): 182, Buying (total rewards):
 -16.755383069626987, Total Reward: -16.755383069626987
 Target network updated at episode 140
 Episode 142/500, Investment (total actions): 185, Buying (total rewards):
 -14.24353754130425, Total Reward: -14.24353754130425
 Episode 143/500, Investment (total actions): 200, Buying (total rewards):
 -11.659893043892225, Total Reward: -11.659893043892225
 Episode 144/500, Investment (total actions): 186, Buying (total rewards):
 -10.689700216054916, Total Reward: -10.689700216054916

Episode 145/500, Investment (total actions): 169, Buying (total rewards):
 -8.927933970466256, Total Reward: -8.927933970466256
 Episode 146/500, Investment (total actions): 181, Buying (total rewards):
 -13.597975108772516, Total Reward: -13.597975108772516
 Episode 147/500, Investment (total actions): 171, Buying (total rewards):
 -12.956529138609767, Total Reward: -12.956529138609767
 Episode 148/500, Investment (total actions): 182, Buying (total rewards):
 -14.052071465644985, Total Reward: -14.052071465644985
 Episode 149/500, Investment (total actions): 193, Buying (total rewards):
 -16.447289265692234, Total Reward: -16.447289265692234
 Episode 150/500, Investment (total actions): 192, Buying (total rewards):
 -18.91512123681605, Total Reward: -18.91512123681605
 Episode 151/500, Investment (total actions): 185, Buying (total rewards):
 -19.553758619353175, Total Reward: -19.553758619353175
 Target network updated at episode 150
 Episode 152/500, Investment (total actions): 210, Buying (total rewards):
 -19.275569978170097, Total Reward: -19.275569978170097
 Episode 153/500, Investment (total actions): 191, Buying (total rewards):
 -15.790894519537687, Total Reward: -15.790894519537687
 Episode 154/500, Investment (total actions): 179, Buying (total rewards):
 -15.382374168373644, Total Reward: -15.382374168373644
 Episode 155/500, Investment (total actions): 185, Buying (total rewards):
 -15.00438869453501, Total Reward: -15.00438869453501
 Episode 156/500, Investment (total actions): 218, Buying (total rewards):
 -18.236644210584927, Total Reward: -18.236644210584927
 Episode 157/500, Investment (total actions): 199, Buying (total rewards):
 -17.789783666841686, Total Reward: -17.789783666841686
 Episode 158/500, Investment (total actions): 199, Buying (total rewards):
 -15.596179444692098, Total Reward: -15.596179444692098
 Episode 159/500, Investment (total actions): 193, Buying (total rewards):
 -13.473902410827577, Total Reward: -13.473902410827577
 Episode 160/500, Investment (total actions): 206, Buying (total rewards):
 -11.89431238360703, Total Reward: -11.89431238360703
 Episode 161/500, Investment (total actions): 197, Buying (total rewards):
 -10.341887437272817, Total Reward: -10.341887437272817
 Target network updated at episode 160
 Episode 162/500, Investment (total actions): 200, Buying (total rewards):
 -11.576737844152376, Total Reward: -11.576737844152376
 Episode 163/500, Investment (total actions): 191, Buying (total rewards):
 -10.421265911543742, Total Reward: -10.421265911543742
 Episode 164/500, Investment (total actions): 194, Buying (total rewards):
 -11.237614370882511, Total Reward: -11.237614370882511
 Episode 165/500, Investment (total actions): 198, Buying (total rewards):
 -8.973950946703553, Total Reward: -8.973950946703553
 Episode 166/500, Investment (total actions): 203, Buying (total rewards):
 -11.168693130137399, Total Reward: -11.168693130137399
 Episode 167/500, Investment (total actions): 189, Buying (total rewards):
 -9.811738328076899, Total Reward: -9.811738328076899

Episode 168/500, Investment (total actions): 202, Buying (total rewards):
 -8.898543489165604, Total Reward: -8.898543489165604
 Episode 169/500, Investment (total actions): 191, Buying (total rewards):
 -10.013476829393767, Total Reward: -10.013476829393767
 Episode 170/500, Investment (total actions): 180, Buying (total rewards):
 -9.837740617105737, Total Reward: -9.837740617105737
 Episode 171/500, Investment (total actions): 177, Buying (total rewards):
 -9.100561048835516, Total Reward: -9.100561048835516
 Target network updated at episode 170
 Episode 172/500, Investment (total actions): 212, Buying (total rewards):
 -6.453556229360402, Total Reward: -6.453556229360402
 Episode 173/500, Investment (total actions): 175, Buying (total rewards):
 -8.062528483569622, Total Reward: -8.062528483569622
 Episode 174/500, Investment (total actions): 191, Buying (total rewards):
 -9.153560799080878, Total Reward: -9.153560799080878
 Episode 175/500, Investment (total actions): 170, Buying (total rewards):
 -8.988065849989653, Total Reward: -8.988065849989653
 Episode 176/500, Investment (total actions): 200, Buying (total rewards):
 -9.55340604740195, Total Reward: -9.55340604740195
 Episode 177/500, Investment (total actions): 183, Buying (total rewards):
 -12.462930679321289, Total Reward: -12.462930679321289
 Episode 178/500, Investment (total actions): 182, Buying (total rewards):
 -13.440112974203657, Total Reward: -13.440112974203657
 Episode 179/500, Investment (total actions): 204, Buying (total rewards):
 -14.151809441274963, Total Reward: -14.151809441274963
 Episode 180/500, Investment (total actions): 190, Buying (total rewards):
 -13.58155430201441, Total Reward: -13.58155430201441
 Episode 181/500, Investment (total actions): 184, Buying (total rewards):
 -14.507010124623775, Total Reward: -14.507010124623775
 Target network updated at episode 180
 Episode 182/500, Investment (total actions): 194, Buying (total rewards):
 -15.388259278610349, Total Reward: -15.388259278610349
 Episode 183/500, Investment (total actions): 183, Buying (total rewards):
 -15.124745611101389, Total Reward: -15.124745611101389
 Episode 184/500, Investment (total actions): 188, Buying (total rewards):
 -14.882540201302618, Total Reward: -14.882540201302618
 Episode 185/500, Investment (total actions): 187, Buying (total rewards):
 -15.388581969658844, Total Reward: -15.388581969658844
 Episode 186/500, Investment (total actions): 192, Buying (total rewards):
 -15.343352106865495, Total Reward: -15.343352106865495
 Episode 187/500, Investment (total actions): 205, Buying (total rewards):
 -15.967782263178378, Total Reward: -15.967782263178378
 Episode 188/500, Investment (total actions): 201, Buying (total rewards):
 -16.840216051787138, Total Reward: -16.840216051787138
 Episode 189/500, Investment (total actions): 199, Buying (total rewards):
 -18.244874723255634, Total Reward: -18.244874723255634
 Episode 190/500, Investment (total actions): 206, Buying (total rewards):
 -16.33647760655731, Total Reward: -16.33647760655731

Episode 191/500, Investment (total actions): 193, Buying (total rewards):
 -18.028226962313056, Total Reward: -18.028226962313056
 Target network updated at episode 190
 Episode 192/500, Investment (total actions): 187, Buying (total rewards):
 -18.105362420668826, Total Reward: -18.105362420668826
 Episode 193/500, Investment (total actions): 223, Buying (total rewards):
 -18.111946191696916, Total Reward: -18.111946191696916
 Episode 194/500, Investment (total actions): 208, Buying (total rewards):
 -17.919245816767216, Total Reward: -17.919245816767216
 Episode 195/500, Investment (total actions): 189, Buying (total rewards):
 -15.535279083997011, Total Reward: -15.535279083997011
 Episode 196/500, Investment (total actions): 206, Buying (total rewards):
 -17.173184282830334, Total Reward: -17.173184282830334
 Episode 197/500, Investment (total actions): 215, Buying (total rewards):
 -16.728465733584017, Total Reward: -16.728465733584017
 Episode 198/500, Investment (total actions): 201, Buying (total rewards):
 -18.113188231829554, Total Reward: -18.113188231829554
 Episode 199/500, Investment (total actions): 209, Buying (total rewards):
 -18.990975863765925, Total Reward: -18.990975863765925
 Episode 200/500, Investment (total actions): 208, Buying (total rewards):
 -20.404071773402393, Total Reward: -20.404071773402393
 Episode 201/500, Investment (total actions): 213, Buying (total rewards):
 -19.968558862805367, Total Reward: -19.968558862805367
 Target network updated at episode 200
 Episode 202/500, Investment (total actions): 199, Buying (total rewards):
 -20.46309428184759, Total Reward: -20.46309428184759
 Episode 203/500, Investment (total actions): 213, Buying (total rewards):
 -20.14763001073152, Total Reward: -20.14763001073152
 Episode 204/500, Investment (total actions): 214, Buying (total rewards):
 -20.13261380791664, Total Reward: -20.13261380791664
 Episode 205/500, Investment (total actions): 206, Buying (total rewards):
 -18.909862007945776, Total Reward: -18.909862007945776
 Episode 206/500, Investment (total actions): 209, Buying (total rewards):
 -18.9628084897995, Total Reward: -18.9628084897995
 Episode 207/500, Investment (total actions): 192, Buying (total rewards):
 -19.66918515600264, Total Reward: -19.66918515600264
 Episode 208/500, Investment (total actions): 206, Buying (total rewards):
 -18.151904300786555, Total Reward: -18.151904300786555
 Episode 209/500, Investment (total actions): 225, Buying (total rewards):
 -18.285480558755808, Total Reward: -18.285480558755808
 Episode 210/500, Investment (total actions): 218, Buying (total rewards):
 -16.546945463865995, Total Reward: -16.546945463865995
 Episode 211/500, Investment (total actions): 230, Buying (total rewards):
 -14.542238725698553, Total Reward: -14.542238725698553
 Target network updated at episode 210
 Episode 212/500, Investment (total actions): 232, Buying (total rewards):
 -14.373263650108129, Total Reward: -14.373263650108129
 Episode 213/500, Investment (total actions): 224, Buying (total rewards):

-15.919270204380155, Total Reward: -15.919270204380155
 Episode 214/500, Investment (total actions): 221, Buying (total rewards):
 -15.899036803282797, Total Reward: -15.899036803282797
 Episode 215/500, Investment (total actions): 210, Buying (total rewards):
 -13.909969966858625, Total Reward: -13.909969966858625
 Episode 216/500, Investment (total actions): 217, Buying (total rewards):
 -15.360101342666894, Total Reward: -15.360101342666894
 Episode 217/500, Investment (total actions): 210, Buying (total rewards):
 -16.455330468714237, Total Reward: -16.455330468714237
 Episode 218/500, Investment (total actions): 236, Buying (total rewards):
 -15.985434942616848, Total Reward: -15.985434942616848
 Episode 219/500, Investment (total actions): 232, Buying (total rewards):
 -16.857911225408316, Total Reward: -16.857911225408316
 Episode 220/500, Investment (total actions): 216, Buying (total rewards):
 -17.070005420362577, Total Reward: -17.070005420362577
 Episode 221/500, Investment (total actions): 209, Buying (total rewards):
 -16.840769258444197, Total Reward: -16.840769258444197
 Target network updated at episode 220
 Episode 222/500, Investment (total actions): 211, Buying (total rewards):
 -16.79500823142007, Total Reward: -16.79500823142007
 Episode 223/500, Investment (total actions): 204, Buying (total rewards):
 -17.470620850566775, Total Reward: -17.470620850566775
 Episode 224/500, Investment (total actions): 203, Buying (total rewards):
 -17.939579546335153, Total Reward: -17.939579546335153
 Episode 225/500, Investment (total actions): 202, Buying (total rewards):
 -17.51490018237382, Total Reward: -17.51490018237382
 Episode 226/500, Investment (total actions): 204, Buying (total rewards):
 -16.449684188701212, Total Reward: -16.449684188701212
 Episode 227/500, Investment (total actions): 199, Buying (total rewards):
 -17.860403028782457, Total Reward: -17.860403028782457
 Episode 228/500, Investment (total actions): 207, Buying (total rewards):
 -18.63236552104354, Total Reward: -18.63236552104354
 Episode 229/500, Investment (total actions): 201, Buying (total rewards):
 -18.213800760917366, Total Reward: -18.213800760917366
 Episode 230/500, Investment (total actions): 209, Buying (total rewards):
 -17.623339843004942, Total Reward: -17.623339843004942
 Episode 231/500, Investment (total actions): 214, Buying (total rewards):
 -17.782208945136517, Total Reward: -17.782208945136517
 Target network updated at episode 230
 Episode 232/500, Investment (total actions): 200, Buying (total rewards):
 -17.15624442510307, Total Reward: -17.15624442510307
 Episode 233/500, Investment (total actions): 198, Buying (total rewards):
 -15.185819116421044, Total Reward: -15.185819116421044
 Episode 234/500, Investment (total actions): 202, Buying (total rewards):
 -15.669458623044193, Total Reward: -15.669458623044193
 Episode 235/500, Investment (total actions): 211, Buying (total rewards):
 -15.083403365686536, Total Reward: -15.083403365686536
 Episode 236/500, Investment (total actions): 211, Buying (total rewards):

-14.74111724179238, Total Reward: -14.74111724179238
 Episode 237/500, Investment (total actions): 206, Buying (total rewards):
 -14.372561275959015, Total Reward: -14.372561275959015
 Episode 238/500, Investment (total actions): 210, Buying (total rewards):
 -14.14282364747487, Total Reward: -14.14282364747487
 Episode 239/500, Investment (total actions): 234, Buying (total rewards):
 -12.668649054132402, Total Reward: -12.668649054132402
 Episode 240/500, Investment (total actions): 215, Buying (total rewards):
 -12.486658935435116, Total Reward: -12.486658935435116
 Episode 241/500, Investment (total actions): 202, Buying (total rewards):
 -12.888335756026208, Total Reward: -12.888335756026208
 Target network updated at episode 240
 Episode 242/500, Investment (total actions): 213, Buying (total rewards):
 -12.715538190444931, Total Reward: -12.715538190444931
 Episode 243/500, Investment (total actions): 218, Buying (total rewards):
 -12.412831099238247, Total Reward: -12.412831099238247
 Episode 244/500, Investment (total actions): 194, Buying (total rewards):
 -13.181845373124816, Total Reward: -13.181845373124816
 Episode 245/500, Investment (total actions): 210, Buying (total rewards):
 -13.682733251247555, Total Reward: -13.682733251247555
 Episode 246/500, Investment (total actions): 218, Buying (total rewards):
 -14.198930468061008, Total Reward: -14.198930468061008
 Episode 247/500, Investment (total actions): 239, Buying (total rewards):
 -13.371758076362312, Total Reward: -13.371758076362312
 Episode 248/500, Investment (total actions): 212, Buying (total rewards):
 -12.782658288720995, Total Reward: -12.782658288720995
 Episode 249/500, Investment (total actions): 228, Buying (total rewards):
 -11.930339299608022, Total Reward: -11.930339299608022
 Episode 250/500, Investment (total actions): 213, Buying (total rewards):
 -11.246064975392073, Total Reward: -11.246064975392073
 Episode 251/500, Investment (total actions): 199, Buying (total rewards):
 -9.944390491582453, Total Reward: -9.944390491582453
 Target network updated at episode 250
 Episode 252/500, Investment (total actions): 187, Buying (total rewards):
 -10.9162967344746, Total Reward: -10.9162967344746
 Episode 253/500, Investment (total actions): 200, Buying (total rewards):
 -10.98434423469007, Total Reward: -10.98434423469007
 Episode 254/500, Investment (total actions): 214, Buying (total rewards):
 -10.964435077738017, Total Reward: -10.964435077738017
 Episode 255/500, Investment (total actions): 195, Buying (total rewards):
 -10.423149621114135, Total Reward: -10.423149621114135
 Episode 256/500, Investment (total actions): 212, Buying (total rewards):
 -10.83054561726749, Total Reward: -10.83054561726749
 Episode 257/500, Investment (total actions): 200, Buying (total rewards):
 -10.928145945072174, Total Reward: -10.928145945072174
 Episode 258/500, Investment (total actions): 205, Buying (total rewards):
 -11.645673721912317, Total Reward: -11.645673721912317
 Episode 259/500, Investment (total actions): 202, Buying (total rewards):

-10.725445751100779, Total Reward: -10.725445751100779
 Episode 260/500, Investment (total actions): 206, Buying (total rewards):
 -11.01796093210578, Total Reward: -11.01796093210578
 Episode 261/500, Investment (total actions): 201, Buying (total rewards):
 -10.606105971732177, Total Reward: -10.606105971732177
 Target network updated at episode 260
 Episode 262/500, Investment (total actions): 196, Buying (total rewards):
 -10.409426721860655, Total Reward: -10.409426721860655
 Episode 263/500, Investment (total actions): 199, Buying (total rewards):
 -10.478567499667406, Total Reward: -10.478567499667406
 Episode 264/500, Investment (total actions): 194, Buying (total rewards):
 -10.821823596954346, Total Reward: -10.821823596954346
 Episode 265/500, Investment (total actions): 209, Buying (total rewards):
 -10.974714914569631, Total Reward: -10.974714914569631
 Episode 266/500, Investment (total actions): 197, Buying (total rewards):
 -11.502299519255757, Total Reward: -11.502299519255757
 Episode 267/500, Investment (total actions): 208, Buying (total rewards):
 -11.944427892100066, Total Reward: -11.944427892100066
 Episode 268/500, Investment (total actions): 202, Buying (total rewards):
 -12.978518379153684, Total Reward: -12.978518379153684
 Episode 269/500, Investment (total actions): 211, Buying (total rewards):
 -12.518906399607658, Total Reward: -12.518906399607658
 Episode 270/500, Investment (total actions): 207, Buying (total rewards):
 -13.012090435251594, Total Reward: -13.012090435251594
 Episode 271/500, Investment (total actions): 198, Buying (total rewards):
 -12.368564395233989, Total Reward: -12.368564395233989
 Target network updated at episode 270
 Episode 272/500, Investment (total actions): 200, Buying (total rewards):
 -12.338380155153573, Total Reward: -12.338380155153573
 Episode 273/500, Investment (total actions): 186, Buying (total rewards):
 -12.583246232476085, Total Reward: -12.583246232476085
 Episode 274/500, Investment (total actions): 202, Buying (total rewards):
 -12.411163602024317, Total Reward: -12.411163602024317
 Episode 275/500, Investment (total actions): 197, Buying (total rewards):
 -12.012400043196976, Total Reward: -12.012400043196976
 Episode 276/500, Investment (total actions): 199, Buying (total rewards):
 -12.082343221176416, Total Reward: -12.082343221176416
 Episode 277/500, Investment (total actions): 204, Buying (total rewards):
 -12.382675924221985, Total Reward: -12.382675924221985
 Episode 278/500, Investment (total actions): 225, Buying (total rewards):
 -12.459012857405469, Total Reward: -12.459012857405469
 Episode 279/500, Investment (total actions): 215, Buying (total rewards):
 -11.944242917001247, Total Reward: -11.944242917001247
 Episode 280/500, Investment (total actions): 199, Buying (total rewards):
 -11.805422956123948, Total Reward: -11.805422956123948
 Episode 281/500, Investment (total actions): 200, Buying (total rewards):
 -12.06105968914926, Total Reward: -12.06105968914926
 Target network updated at episode 280

Episode 282/500, Investment (total actions): 196, Buying (total rewards):
 -12.12182847186341, Total Reward: -12.12182847186341
 Episode 283/500, Investment (total actions): 208, Buying (total rewards):
 -11.772173080593348, Total Reward: -11.772173080593348
 Episode 284/500, Investment (total actions): 210, Buying (total rewards):
 -12.087448611389846, Total Reward: -12.087448611389846
 Episode 285/500, Investment (total actions): 213, Buying (total rewards):
 -11.929606214631349, Total Reward: -11.929606214631349
 Episode 286/500, Investment (total actions): 217, Buying (total rewards):
 -12.090513482224196, Total Reward: -12.090513482224196
 Episode 287/500, Investment (total actions): 211, Buying (total rewards):
 -12.295293759554625, Total Reward: -12.295293759554625
 Episode 288/500, Investment (total actions): 215, Buying (total rewards):
 -12.292322964407504, Total Reward: -12.292322964407504
 Episode 289/500, Investment (total actions): 209, Buying (total rewards):
 -11.57120988261886, Total Reward: -11.57120988261886
 Episode 290/500, Investment (total actions): 217, Buying (total rewards):
 -11.597606416791677, Total Reward: -11.597606416791677
 Episode 291/500, Investment (total actions): 223, Buying (total rewards):
 -11.373556697159074, Total Reward: -11.373556697159074
 Target network updated at episode 290
 Episode 292/500, Investment (total actions): 225, Buying (total rewards):
 -11.447111553512514, Total Reward: -11.447111553512514
 Episode 293/500, Investment (total actions): 226, Buying (total rewards):
 -10.985841805115342, Total Reward: -10.985841805115342
 Episode 294/500, Investment (total actions): 222, Buying (total rewards):
 -11.44238394824788, Total Reward: -11.44238394824788
 Episode 295/500, Investment (total actions): 231, Buying (total rewards):
 -10.91081125382334, Total Reward: -10.91081125382334
 Episode 296/500, Investment (total actions): 238, Buying (total rewards):
 -11.303054826334119, Total Reward: -11.303054826334119
 Episode 297/500, Investment (total actions): 234, Buying (total rewards):
 -10.346089554950595, Total Reward: -10.346089554950595
 Episode 298/500, Investment (total actions): 223, Buying (total rewards):
 -10.506319700740278, Total Reward: -10.506319700740278
 Episode 299/500, Investment (total actions): 223, Buying (total rewards):
 -10.1510204449296, Total Reward: -10.1510204449296
 Episode 300/500, Investment (total actions): 233, Buying (total rewards):
 -10.340568786486983, Total Reward: -10.340568786486983
 Episode 301/500, Investment (total actions): 230, Buying (total rewards):
 -10.018741535954177, Total Reward: -10.018741535954177
 Target network updated at episode 300
 Episode 302/500, Investment (total actions): 243, Buying (total rewards):
 -9.548941296539851, Total Reward: -9.548941296539851
 Episode 303/500, Investment (total actions): 223, Buying (total rewards):
 -8.835550513816997, Total Reward: -8.835550513816997
 Episode 304/500, Investment (total actions): 234, Buying (total rewards):
 -8.027823815820739, Total Reward: -8.027823815820739

Episode 305/500, Investment (total actions): 224, Buying (total rewards):
 -7.1455072639510036, Total Reward: -7.1455072639510036
 Episode 306/500, Investment (total actions): 238, Buying (total rewards):
 -6.986908186576329, Total Reward: -6.986908186576329
 Episode 307/500, Investment (total actions): 238, Buying (total rewards):
 -7.206466569332406, Total Reward: -7.206466569332406
 Episode 308/500, Investment (total actions): 222, Buying (total rewards):
 -7.264536146540195, Total Reward: -7.264536146540195
 Episode 309/500, Investment (total actions): 227, Buying (total rewards):
 -7.093166980368551, Total Reward: -7.093166980368551
 Episode 310/500, Investment (total actions): 216, Buying (total rewards):
 -6.49506825953722, Total Reward: -6.49506825953722
 Episode 311/500, Investment (total actions): 215, Buying (total rewards):
 -5.59352065785788, Total Reward: -5.59352065785788
 Target network updated at episode 310
 Episode 312/500, Investment (total actions): 214, Buying (total rewards):
 -5.966402786783874, Total Reward: -5.966402786783874
 Episode 313/500, Investment (total actions): 216, Buying (total rewards):
 -5.991622281260788, Total Reward: -5.991622281260788
 Episode 314/500, Investment (total actions): 208, Buying (total rewards):
 -5.469410019926727, Total Reward: -5.469410019926727
 Episode 315/500, Investment (total actions): 235, Buying (total rewards):
 -5.19254011567682, Total Reward: -5.19254011567682
 Episode 316/500, Investment (total actions): 231, Buying (total rewards):
 -4.839510613237508, Total Reward: -4.839510613237508
 Episode 317/500, Investment (total actions): 230, Buying (total rewards):
 -4.2403313079848886, Total Reward: -4.2403313079848886
 Episode 318/500, Investment (total actions): 229, Buying (total rewards):
 -4.632245515473187, Total Reward: -4.632245515473187
 Episode 319/500, Investment (total actions): 228, Buying (total rewards):
 -4.4965744013898075, Total Reward: -4.4965744013898075
 Episode 320/500, Investment (total actions): 223, Buying (total rewards):
 -4.1403231819858775, Total Reward: -4.1403231819858775
 Episode 321/500, Investment (total actions): 229, Buying (total rewards):
 -4.4828251213766634, Total Reward: -4.4828251213766634
 Target network updated at episode 320
 Episode 322/500, Investment (total actions): 217, Buying (total rewards):
 -4.02746033994481, Total Reward: -4.02746033994481
 Episode 323/500, Investment (total actions): 227, Buying (total rewards):
 -3.7053970429114997, Total Reward: -3.7053970429114997
 Episode 324/500, Investment (total actions): 220, Buying (total rewards):
 -3.6372114438563585, Total Reward: -3.6372114438563585
 Episode 325/500, Investment (total actions): 219, Buying (total rewards):
 -3.4876498159137554, Total Reward: -3.4876498159137554
 Episode 326/500, Investment (total actions): 223, Buying (total rewards):
 -3.444570926949382, Total Reward: -3.444570926949382
 Episode 327/500, Investment (total actions): 231, Buying (total rewards):
 -3.0548424927983433, Total Reward: -3.0548424927983433

Episode 328/500, Investment (total actions): 232, Buying (total rewards):
 -3.141789418645203, Total Reward: -3.141789418645203
 Episode 329/500, Investment (total actions): 230, Buying (total rewards):
 -3.296841736184433, Total Reward: -3.296841736184433
 Episode 330/500, Investment (total actions): 220, Buying (total rewards):
 -3.5360784417716786, Total Reward: -3.5360784417716786
 Episode 331/500, Investment (total actions): 218, Buying (total rewards):
 -3.7188303992152214, Total Reward: -3.7188303992152214
 Target network updated at episode 330
 Episode 332/500, Investment (total actions): 228, Buying (total rewards):
 -3.834081714041531, Total Reward: -3.834081714041531
 Episode 333/500, Investment (total actions): 210, Buying (total rewards):
 -3.641167657915503, Total Reward: -3.641167657915503
 Episode 334/500, Investment (total actions): 229, Buying (total rewards):
 -3.230866447673179, Total Reward: -3.230866447673179
 Episode 335/500, Investment (total actions): 227, Buying (total rewards):
 -3.063870548387058, Total Reward: -3.063870548387058
 Episode 336/500, Investment (total actions): 224, Buying (total rewards):
 -3.3115713309962302, Total Reward: -3.3115713309962302
 Episode 337/500, Investment (total actions): 202, Buying (total rewards):
 -3.3131010830402374, Total Reward: -3.3131010830402374
 Episode 338/500, Investment (total actions): 213, Buying (total rewards):
 -3.894079852849245, Total Reward: -3.894079852849245
 Episode 339/500, Investment (total actions): 211, Buying (total rewards):
 -4.122069815173745, Total Reward: -4.122069815173745
 Episode 340/500, Investment (total actions): 217, Buying (total rewards):
 -4.038618795573711, Total Reward: -4.038618795573711
 Episode 341/500, Investment (total actions): 209, Buying (total rewards):
 -4.217239119112492, Total Reward: -4.217239119112492
 Target network updated at episode 340
 Episode 342/500, Investment (total actions): 202, Buying (total rewards):
 -4.727823602035642, Total Reward: -4.727823602035642
 Episode 343/500, Investment (total actions): 214, Buying (total rewards):
 -4.960545163601637, Total Reward: -4.960545163601637
 Episode 344/500, Investment (total actions): 210, Buying (total rewards):
 -4.7501056380569935, Total Reward: -4.7501056380569935
 Episode 345/500, Investment (total actions): 200, Buying (total rewards):
 -4.755855293478817, Total Reward: -4.755855293478817
 Episode 346/500, Investment (total actions): 205, Buying (total rewards):
 -5.010019784560427, Total Reward: -5.010019784560427
 Episode 347/500, Investment (total actions): 202, Buying (total rewards):
 -4.862878271378577, Total Reward: -4.862878271378577
 Episode 348/500, Investment (total actions): 210, Buying (total rewards):
 -5.002865309535991, Total Reward: -5.002865309535991
 Episode 349/500, Investment (total actions): 197, Buying (total rewards):
 -5.218791755381972, Total Reward: -5.218791755381972
 Episode 350/500, Investment (total actions): 195, Buying (total rewards):
 -5.5559450522996485, Total Reward: -5.5559450522996485

Episode 351/500, Investment (total actions): 207, Buying (total rewards):
 -5.4019907880574465, Total Reward: -5.4019907880574465
 Target network updated at episode 350
 Episode 352/500, Investment (total actions): 207, Buying (total rewards):
 -5.320374886505306, Total Reward: -5.320374886505306
 Episode 353/500, Investment (total actions): 207, Buying (total rewards):
 -5.171941302192863, Total Reward: -5.171941302192863
 Episode 354/500, Investment (total actions): 204, Buying (total rewards):
 -5.148933213204145, Total Reward: -5.148933213204145
 Episode 355/500, Investment (total actions): 202, Buying (total rewards):
 -4.827212918084115, Total Reward: -4.827212918084115
 Episode 356/500, Investment (total actions): 220, Buying (total rewards):
 -5.259326248138677, Total Reward: -5.259326248138677
 Episode 357/500, Investment (total actions): 219, Buying (total rewards):
 -5.437755957245827, Total Reward: -5.437755957245827
 Episode 358/500, Investment (total actions): 216, Buying (total rewards):
 -5.611555737443268, Total Reward: -5.611555737443268
 Episode 359/500, Investment (total actions): 204, Buying (total rewards):
 -5.9850418251007795, Total Reward: -5.9850418251007795
 Episode 360/500, Investment (total actions): 212, Buying (total rewards):
 -6.129153661429882, Total Reward: -6.129153661429882
 Episode 361/500, Investment (total actions): 215, Buying (total rewards):
 -6.101569151505828, Total Reward: -6.101569151505828
 Target network updated at episode 360
 Episode 362/500, Investment (total actions): 213, Buying (total rewards):
 -6.035208523273468, Total Reward: -6.035208523273468
 Episode 363/500, Investment (total actions): 216, Buying (total rewards):
 -6.151057488794322, Total Reward: -6.151057488794322
 Episode 364/500, Investment (total actions): 219, Buying (total rewards):
 -6.35354204941541, Total Reward: -6.35354204941541
 Episode 365/500, Investment (total actions): 208, Buying (total rewards):
 -6.3065500957891345, Total Reward: -6.3065500957891345
 Episode 366/500, Investment (total actions): 224, Buying (total rewards):
 -6.388560477644205, Total Reward: -6.388560477644205
 Episode 367/500, Investment (total actions): 224, Buying (total rewards):
 -6.742798593826592, Total Reward: -6.742798593826592
 Episode 368/500, Investment (total actions): 218, Buying (total rewards):
 -6.978838445618749, Total Reward: -6.978838445618749
 Episode 369/500, Investment (total actions): 222, Buying (total rewards):
 -7.085321792401373, Total Reward: -7.085321792401373
 Episode 370/500, Investment (total actions): 222, Buying (total rewards):
 -7.130264946259558, Total Reward: -7.130264946259558
 Episode 371/500, Investment (total actions): 222, Buying (total rewards):
 -7.382972361636348, Total Reward: -7.382972361636348
 Target network updated at episode 370
 Episode 372/500, Investment (total actions): 215, Buying (total rewards):
 -7.8635367741808295, Total Reward: -7.8635367741808295
 Episode 373/500, Investment (total actions): 224, Buying (total rewards):

-7.663057930767536, Total Reward: -7.663057930767536
 Episode 374/500, Investment (total actions): 222, Buying (total rewards):
 -7.559992938302457, Total Reward: -7.559992938302457
 Episode 375/500, Investment (total actions): 216, Buying (total rewards):
 -7.758984582033008, Total Reward: -7.758984582033008
 Episode 376/500, Investment (total actions): 229, Buying (total rewards):
 -7.792058388818987, Total Reward: -7.792058388818987
 Episode 377/500, Investment (total actions): 230, Buying (total rewards):
 -7.678511378355324, Total Reward: -7.678511378355324
 Episode 378/500, Investment (total actions): 214, Buying (total rewards):
 -7.4477173974737525, Total Reward: -7.4477173974737525
 Episode 379/500, Investment (total actions): 222, Buying (total rewards):
 -7.508300217567012, Total Reward: -7.508300217567012
 Episode 380/500, Investment (total actions): 221, Buying (total rewards):
 -7.4500637403689325, Total Reward: -7.4500637403689325
 Episode 381/500, Investment (total actions): 219, Buying (total rewards):
 -7.5943829622119665, Total Reward: -7.5943829622119665
 Target network updated at episode 380
 Episode 382/500, Investment (total actions): 213, Buying (total rewards):
 -7.341388339176774, Total Reward: -7.341388339176774
 Episode 383/500, Investment (total actions): 219, Buying (total rewards):
 -7.481537878513336, Total Reward: -7.481537878513336
 Episode 384/500, Investment (total actions): 225, Buying (total rewards):
 -7.367173161357641, Total Reward: -7.367173161357641
 Episode 385/500, Investment (total actions): 224, Buying (total rewards):
 -7.253116929903626, Total Reward: -7.253116929903626
 Episode 386/500, Investment (total actions): 224, Buying (total rewards):
 -7.17136799171567, Total Reward: -7.17136799171567
 Episode 387/500, Investment (total actions): 215, Buying (total rewards):
 -7.007459435611963, Total Reward: -7.007459435611963
 Episode 388/500, Investment (total actions): 216, Buying (total rewards):
 -6.726338624954224, Total Reward: -6.726338624954224
 Episode 389/500, Investment (total actions): 218, Buying (total rewards):
 -6.59129395801574, Total Reward: -6.59129395801574
 Episode 390/500, Investment (total actions): 209, Buying (total rewards):
 -6.28298981487751, Total Reward: -6.28298981487751
 Episode 391/500, Investment (total actions): 207, Buying (total rewards):
 -6.137095883022994, Total Reward: -6.137095883022994
 Target network updated at episode 390
 Episode 392/500, Investment (total actions): 220, Buying (total rewards):
 -6.507127804681659, Total Reward: -6.507127804681659
 Episode 393/500, Investment (total actions): 219, Buying (total rewards):
 -6.272584740072489, Total Reward: -6.272584740072489
 Episode 394/500, Investment (total actions): 226, Buying (total rewards):
 -6.221612895024009, Total Reward: -6.221612895024009
 Episode 395/500, Investment (total actions): 221, Buying (total rewards):
 -6.2875848123803735, Total Reward: -6.2875848123803735
 Episode 396/500, Investment (total actions): 223, Buying (total rewards):

-6.5771135311806574, Total Reward: -6.5771135311806574
 Episode 397/500, Investment (total actions): 229, Buying (total rewards):
 -6.808737627347, Total Reward: -6.808737627347
 Episode 398/500, Investment (total actions): 220, Buying (total rewards):
 -7.015933339949697, Total Reward: -7.015933339949697
 Episode 399/500, Investment (total actions): 225, Buying (total rewards):
 -7.230961409482916, Total Reward: -7.230961409482916
 Episode 400/500, Investment (total actions): 219, Buying (total rewards):
 -7.42672484042123, Total Reward: -7.42672484042123
 Episode 401/500, Investment (total actions): 213, Buying (total rewards):
 -7.321915050968528, Total Reward: -7.321915050968528
 Target network updated at episode 400
 Episode 402/500, Investment (total actions): 219, Buying (total rewards):
 -7.376588340848684, Total Reward: -7.376588340848684
 Episode 403/500, Investment (total actions): 227, Buying (total rewards):
 -7.515005061402917, Total Reward: -7.515005061402917
 Episode 404/500, Investment (total actions): 223, Buying (total rewards):
 -7.63634317368269, Total Reward: -7.63634317368269
 Episode 405/500, Investment (total actions): 224, Buying (total rewards):
 -7.702447765972465, Total Reward: -7.702447765972465
 Episode 406/500, Investment (total actions): 220, Buying (total rewards):
 -7.76642574602738, Total Reward: -7.76642574602738
 Episode 407/500, Investment (total actions): 214, Buying (total rewards):
 -7.665894993580878, Total Reward: -7.665894993580878
 Episode 408/500, Investment (total actions): 212, Buying (total rewards):
 -7.732213805196807, Total Reward: -7.732213805196807
 Episode 409/500, Investment (total actions): 209, Buying (total rewards):
 -7.832913927966729, Total Reward: -7.832913927966729
 Episode 410/500, Investment (total actions): 210, Buying (total rewards):
 -7.884399672970176, Total Reward: -7.884399672970176
 Episode 411/500, Investment (total actions): 221, Buying (total rewards):
 -7.913129759021103, Total Reward: -7.913129759021103
 Target network updated at episode 410
 Episode 412/500, Investment (total actions): 224, Buying (total rewards):
 -8.044021800160408, Total Reward: -8.044021800160408
 Episode 413/500, Investment (total actions): 209, Buying (total rewards):
 -7.923264101380482, Total Reward: -7.923264101380482
 Episode 414/500, Investment (total actions): 219, Buying (total rewards):
 -7.846931079402566, Total Reward: -7.846931079402566
 Episode 415/500, Investment (total actions): 229, Buying (total rewards):
 -7.332476703450084, Total Reward: -7.332476703450084
 Episode 416/500, Investment (total actions): 217, Buying (total rewards):
 -6.914663138799369, Total Reward: -6.914663138799369
 Episode 417/500, Investment (total actions): 228, Buying (total rewards):
 -6.858390711146058, Total Reward: -6.858390711146058
 Episode 418/500, Investment (total actions): 227, Buying (total rewards):
 -6.777209305670112, Total Reward: -6.777209305670112
 Episode 419/500, Investment (total actions): 229, Buying (total rewards):

-6.73368520475924, Total Reward: -6.73368520475924
 Episode 420/500, Investment (total actions): 227, Buying (total rewards):
 -6.610830704914406, Total Reward: -6.610830704914406
 Episode 421/500, Investment (total actions): 216, Buying (total rewards):
 -6.705346715636551, Total Reward: -6.705346715636551
 Target network updated at episode 420
 Episode 422/500, Investment (total actions): 228, Buying (total rewards):
 -6.925367619842291, Total Reward: -6.925367619842291
 Episode 423/500, Investment (total actions): 215, Buying (total rewards):
 -6.923600754700601, Total Reward: -6.923600754700601
 Episode 424/500, Investment (total actions): 217, Buying (total rewards):
 -6.822948012501001, Total Reward: -6.822948012501001
 Episode 425/500, Investment (total actions): 226, Buying (total rewards):
 -6.7611622251570225, Total Reward: -6.7611622251570225
 Episode 426/500, Investment (total actions): 224, Buying (total rewards):
 -6.605216217925772, Total Reward: -6.605216217925772
 Episode 427/500, Investment (total actions): 226, Buying (total rewards):
 -6.066176884691231, Total Reward: -6.066176884691231
 Episode 428/500, Investment (total actions): 224, Buying (total rewards):
 -5.970580950845033, Total Reward: -5.970580950845033
 Episode 429/500, Investment (total actions): 226, Buying (total rewards):
 -5.805578103289008, Total Reward: -5.805578103289008
 Episode 430/500, Investment (total actions): 215, Buying (total rewards):
 -5.614888512995094, Total Reward: -5.614888512995094
 Episode 431/500, Investment (total actions): 230, Buying (total rewards):
 -5.736274645430967, Total Reward: -5.736274645430967
 Target network updated at episode 430
 Episode 432/500, Investment (total actions): 236, Buying (total rewards):
 -5.675538022071123, Total Reward: -5.675538022071123
 Episode 433/500, Investment (total actions): 237, Buying (total rewards):
 -5.789553066715598, Total Reward: -5.789553066715598
 Episode 434/500, Investment (total actions): 232, Buying (total rewards):
 -5.920991166494787, Total Reward: -5.920991166494787
 Episode 435/500, Investment (total actions): 237, Buying (total rewards):
 -5.973151371348649, Total Reward: -5.973151371348649
 Episode 436/500, Investment (total actions): 236, Buying (total rewards):
 -6.078697793185711, Total Reward: -6.078697793185711
 Episode 437/500, Investment (total actions): 229, Buying (total rewards):
 -6.104980996111408, Total Reward: -6.104980996111408
 Episode 438/500, Investment (total actions): 237, Buying (total rewards):
 -6.210324025538284, Total Reward: -6.210324025538284
 Episode 439/500, Investment (total actions): 234, Buying (total rewards):
 -6.328389172907919, Total Reward: -6.328389172907919
 Episode 440/500, Investment (total actions): 240, Buying (total rewards):
 -6.344125690869987, Total Reward: -6.344125690869987
 Episode 441/500, Investment (total actions): 238, Buying (total rewards):
 -6.54485918226419, Total Reward: -6.54485918226419
 Target network updated at episode 440

Episode 442/500, Investment (total actions): 240, Buying (total rewards):
 -6.6270557646639645, Total Reward: -6.6270557646639645
 Episode 443/500, Investment (total actions): 233, Buying (total rewards):
 -6.510576338274404, Total Reward: -6.510576338274404
 Episode 444/500, Investment (total actions): 232, Buying (total rewards):
 -6.378462817519903, Total Reward: -6.378462817519903
 Episode 445/500, Investment (total actions): 230, Buying (total rewards):
 -6.607942116446793, Total Reward: -6.607942116446793
 Episode 446/500, Investment (total actions): 233, Buying (total rewards):
 -6.780609641224146, Total Reward: -6.780609641224146
 Episode 447/500, Investment (total actions): 223, Buying (total rewards):
 -6.592011645436287, Total Reward: -6.592011645436287
 Episode 448/500, Investment (total actions): 238, Buying (total rewards):
 -6.609434712678194, Total Reward: -6.609434712678194
 Episode 449/500, Investment (total actions): 231, Buying (total rewards):
 -6.626635329797864, Total Reward: -6.626635329797864
 Episode 450/500, Investment (total actions): 228, Buying (total rewards):
 -6.49082151427865, Total Reward: -6.49082151427865
 Episode 451/500, Investment (total actions): 230, Buying (total rewards):
 -6.306371264159679, Total Reward: -6.306371264159679
 Target network updated at episode 450
 Episode 452/500, Investment (total actions): 226, Buying (total rewards):
 -6.257108716061339, Total Reward: -6.257108716061339
 Episode 453/500, Investment (total actions): 220, Buying (total rewards):
 -6.261895616771653, Total Reward: -6.261895616771653
 Episode 454/500, Investment (total actions): 230, Buying (total rewards):
 -6.203639545012265, Total Reward: -6.203639545012265
 Episode 455/500, Investment (total actions): 220, Buying (total rewards):
 -6.113702218979597, Total Reward: -6.113702218979597
 Episode 456/500, Investment (total actions): 219, Buying (total rewards):
 -6.002204006537795, Total Reward: -6.002204006537795
 Episode 457/500, Investment (total actions): 228, Buying (total rewards):
 -5.913110995054012, Total Reward: -5.913110995054012
 Episode 458/500, Investment (total actions): 222, Buying (total rewards):
 -5.710654346272349, Total Reward: -5.710654346272349
 Episode 459/500, Investment (total actions): 231, Buying (total rewards):
 -5.769690666231327, Total Reward: -5.769690666231327
 Episode 460/500, Investment (total actions): 214, Buying (total rewards):
 -5.766912337392569, Total Reward: -5.766912337392569
 Episode 461/500, Investment (total actions): 219, Buying (total rewards):
 -5.676249640993774, Total Reward: -5.676249640993774
 Target network updated at episode 460
 Episode 462/500, Investment (total actions): 224, Buying (total rewards):
 -5.452003929647617, Total Reward: -5.452003929647617
 Episode 463/500, Investment (total actions): 234, Buying (total rewards):
 -5.374794043600559, Total Reward: -5.374794043600559
 Episode 464/500, Investment (total actions): 222, Buying (total rewards):
 -5.33899780549109, Total Reward: -5.33899780549109

Episode 465/500, Investment (total actions): 231, Buying (total rewards):
 -5.331243811175227, Total Reward: -5.331243811175227
 Episode 466/500, Investment (total actions): 217, Buying (total rewards):
 -5.1062053591012955, Total Reward: -5.1062053591012955
 Episode 467/500, Investment (total actions): 228, Buying (total rewards):
 -5.093719787915688, Total Reward: -5.093719787915688
 Episode 468/500, Investment (total actions): 223, Buying (total rewards):
 -4.895779770100489, Total Reward: -4.895779770100489
 Episode 469/500, Investment (total actions): 231, Buying (total rewards):
 -5.089630675618537, Total Reward: -5.089630675618537
 Episode 470/500, Investment (total actions): 219, Buying (total rewards):
 -4.923549301922321, Total Reward: -4.923549301922321
 Episode 471/500, Investment (total actions): 225, Buying (total rewards):
 -4.734704087517457, Total Reward: -4.734704087517457
 Target network updated at episode 470
 Episode 472/500, Investment (total actions): 217, Buying (total rewards):
 -4.605449363960361, Total Reward: -4.605449363960361
 Episode 473/500, Investment (total actions): 225, Buying (total rewards):
 -4.621620741207153, Total Reward: -4.621620741207153
 Episode 474/500, Investment (total actions): 226, Buying (total rewards):
 -4.5663510747253895, Total Reward: -4.5663510747253895
 Episode 475/500, Investment (total actions): 228, Buying (total rewards):
 -4.492173914331943, Total Reward: -4.492173914331943
 Episode 476/500, Investment (total actions): 216, Buying (total rewards):
 -4.371298605576158, Total Reward: -4.371298605576158
 Episode 477/500, Investment (total actions): 223, Buying (total rewards):
 -4.348457664018497, Total Reward: -4.348457664018497
 Episode 478/500, Investment (total actions): 224, Buying (total rewards):
 -4.315846492885612, Total Reward: -4.315846492885612
 Episode 479/500, Investment (total actions): 222, Buying (total rewards):
 -4.379042464192025, Total Reward: -4.379042464192025
 Episode 480/500, Investment (total actions): 222, Buying (total rewards):
 -4.252536314539611, Total Reward: -4.252536314539611
 Episode 481/500, Investment (total actions): 213, Buying (total rewards):
 -4.174583205487579, Total Reward: -4.174583205487579
 Target network updated at episode 480
 Episode 482/500, Investment (total actions): 220, Buying (total rewards):
 -4.1071500913240016, Total Reward: -4.1071500913240016
 Episode 483/500, Investment (total actions): 218, Buying (total rewards):
 -4.022710980381817, Total Reward: -4.022710980381817
 Episode 484/500, Investment (total actions): 214, Buying (total rewards):
 -3.9042551702586934, Total Reward: -3.9042551702586934
 Episode 485/500, Investment (total actions): 211, Buying (total rewards):
 -3.8798822490498424, Total Reward: -3.8798822490498424
 Episode 486/500, Investment (total actions): 216, Buying (total rewards):
 -3.7278629555366933, Total Reward: -3.7278629555366933
 Episode 487/500, Investment (total actions): 215, Buying (total rewards):
 -3.7214349535061046, Total Reward: -3.7214349535061046

Episode 488/500, Investment (total actions): 215, Buying (total rewards):
 -3.708659012801945, Total Reward: -3.708659012801945
 Episode 489/500, Investment (total actions): 210, Buying (total rewards):
 -3.6346380695467815, Total Reward: -3.6346380695467815
 Episode 490/500, Investment (total actions): 212, Buying (total rewards):
 -3.557821791386232, Total Reward: -3.557821791386232
 Episode 491/500, Investment (total actions): 208, Buying (total rewards):
 -3.4297507852315903, Total Reward: -3.4297507852315903
 Target network updated at episode 490
 Episode 492/500, Investment (total actions): 213, Buying (total rewards):
 -3.2157176546752453, Total Reward: -3.2157176546752453
 Episode 493/500, Investment (total actions): 215, Buying (total rewards):
 -3.1656036879867315, Total Reward: -3.1656036879867315
 Episode 494/500, Investment (total actions): 214, Buying (total rewards):
 -3.075444532558322, Total Reward: -3.075444532558322
 Episode 495/500, Investment (total actions): 220, Buying (total rewards):
 -3.050139804603532, Total Reward: -3.050139804603532
 Episode 496/500, Investment (total actions): 212, Buying (total rewards):
 -3.0564677622169256, Total Reward: -3.0564677622169256
 Episode 497/500, Investment (total actions): 218, Buying (total rewards):
 -3.101286491379142, Total Reward: -3.101286491379142
 Episode 498/500, Investment (total actions): 209, Buying (total rewards):
 -3.0897391848266125, Total Reward: -3.0897391848266125
 Episode 499/500, Investment (total actions): 216, Buying (total rewards):
 -3.0291587989777327, Total Reward: -3.0291587989777327
 Episode 500/500, Investment (total actions): 215, Buying (total rewards):
 -3.0811828644946218, Total Reward: -3.0811828644946218
 DQN Training completed!

1.0.9 Design network DDQN (0.5 M)

Designing the DDQN model with fully connected layers same as DQN model

```

[16]: ### write your code below this line
      # Define the DDQN architecture (same as DQN)
      class DQN(nn.Module):
          def __init__(self, state_size, action_size):
              super(DQN, self).__init__()
              self.fc1 = nn.Linear(state_size, 64) # First fully connected layer
              self.fc2 = nn.Linear(64, 128) # Second fully connected layer
              self.fc3 = nn.Linear(128, action_size) # Output layer with number of
              ↪ actions

          def forward(self, x):
              x = F.relu(self.fc1(x)) # Apply ReLU activation after the first layer
              x = F.relu(self.fc2(x)) # Apply ReLU activation after the second layer
              return self.fc3(x) # Output action values
  
```

```

# Initialize device to use GPU if available
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

# Initialize Q-network and Target network for DDQN
ddqn_q_network = DQN(env.get_state_space_size(), env.get_action_space_size()).
    ↪to(device)
ddqn_target_network = DQN(env.get_state_space_size(), env.
    ↪get_action_space_size()).to(device)

# Copy the weights from Q-network to Target network (initially they are the_
    ↪same)
ddqn_target_network.load_state_dict(ddqn_q_network.state_dict())
ddqn_target_network.eval() # Target network is in evaluation mode (not updated_
    ↪during training)

# Optimizer for the Q-network (only the Q-network is trained)
optimizer = optim.Adam(ddqn_q_network.parameters(), lr=learning_rate)

# Loss function (Mean Squared Error)
loss_fn = nn.MSELoss()

# Optional: Print model summary
print("DDQN Q-Network Summary:")
summary(ddqn_q_network, input_size=(batch_size, env.get_state_space_size()))

print("DDQN Target Network Summary:")
summary(ddqn_target_network, input_size=(batch_size, env.
    ↪get_state_space_size()))

```

DDQN Q-Network Summary:

DDQN Target Network Summary:

```

[16]: =====
=====
Layer (type:depth-idx)          Output Shape          Param #
=====
=====
DQN                             [64, 3]              --
  Linear: 1-1                   [64, 64]             1,472
  Linear: 1-2                   [64, 128]            8,320
  Linear: 1-3                   [64, 3]              387
=====
=====
Total params: 10,179
Trainable params: 10,179
Non-trainable params: 0
Total mult-adds (M): 0.65

```

```
=====
=====
Input size (MB): 0.01
Forward/backward pass size (MB): 0.10
Params size (MB): 0.04
Estimated Total Size (MB): 0.15
=====
=====
```

1.0.10 Training Iterations (0.5 M)

Note: print all the episodes with the values of investment and buying. (if not printed then -1 will be done.)

```
[17]: ### write your code below this line

# Initialize counters for actions for both DQN and DDQN
ddqn_action_counts = {'decrease': 0, 'maintain': 0, 'increase': 0}

# DDQN Training Loop
def train_ddqn(env, num_episodes=500, batch_size=64, gamma=0.99,
    ↪epsilon_start=1.0, epsilon_end=0.1, epsilon_decay=0.995, learning_rate=0.
    ↪001, target_update_period=100):
    device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
    ddqn_q_network = DQN(env.get_state_space_size(), env.
    ↪get_action_space_size()).to(device)
    ddqn_target_network = DQN(env.get_state_space_size(), env.
    ↪get_action_space_size()).to(device)
    ddqn_target_network.load_state_dict(ddqn_q_network.state_dict())

    optimizer = optim.Adam(ddqn_q_network.parameters(), lr=learning_rate)
    loss_fn = nn.MSELoss()

    replay_buffer = PrioritizedReplayBuffer(max_size=max_replay_memory)
    ddqn_rewards = []
    epsilon = epsilon_start

    # Iterate through number of episodes defined
    for episode in range(num_episodes):
        state = env.reset()
        state = torch.FloatTensor(state).to(device).unsqueeze(0)
        total_reward = 0
        total_investment = 0 # Track total investment
        total_buying = 0 # Track total rewards
        done = False
        step_count = 0
```

```

while not done:
    # Epsilon-greedy action selection
    if np.random.rand() < epsilon:
        action = env.action_space_sample()
    else:
        with torch.no_grad():
            q_values = ddqn_q_network(state)
            action = torch.argmax(q_values).item()

    # Track action counts for DDQN
    if action == 0:
        ddqn_action_counts['decrease'] += 1
    elif action == 1:
        ddqn_action_counts['maintain'] += 1
    else:
        ddqn_action_counts['increase'] += 1

    next_state, reward, done = env.step(action)
    next_state = torch.FloatTensor(next_state).to(device).unsqueeze(0)
    reward = torch.tensor([reward], dtype=torch.float32, device=device)

    # Track investment and buying
    total_investment += action # Assuming action represents energy
    ↪adjustment
    total_buying += reward.item() # Reward represents cost or benefit

    # Store experience in replay buffer
    replay_buffer.add((state, action, reward, next_state, done),
    ↪td_error)

    # Sample from replay buffer and train
    if replay_buffer.size() >= batch_size:
        experiences, indices, weights = replay_buffer.sample(batch_size)
        states, actions, rewards, next_states, dones = zip(*experiences)
        states = torch.cat(states).to(device)
        next_states = torch.cat(next_states).to(device)
        actions = torch.tensor(actions, device=device).unsqueeze(1)
        rewards = torch.cat(rewards).to(device)

        q_values = ddqn_q_network(states).gather(1, actions)
        next_q_values = ddqn_q_network(next_states).max(1)[0].detach()
        target_q_values = ddqn_target_network(next_states).gather(1,
    ↪actions).squeeze()

        targets = rewards + (gamma * target_q_values * (1 - torch.
    ↪tensor(dones, dtype=torch.float32, device=device)))
        loss = loss_fn(q_values.squeeze(), targets)

```



```

        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

        state = next_state
        total_reward += reward.item()
        step_count += 1

    ddqn_rewards.append(total_reward)

    # Decay epsilon
    epsilon = max(epsilon_end, epsilon * epsilon_decay)

    # Update the target network
    if episode % target_update_period == 0:
        ddqn_target_network.load_state_dict(ddqn_q_network.state_dict())

    # Print episode information
    print(f"Episode {episode + 1}/{num_episodes}, Investment (actions):␣
↪{total_investment}, Buying (rewards): {total_buying}, Total Reward:␣
↪{total_reward}")

    # Update the target network every 'target_update_freq' episodes ###
    if episode % target_update_freq == 0:
        target_dqn_model.load_state_dict(dqn_model.state_dict()) # Sync␣
↪the target network weights
        print(f"Target network updated at episode {episode}")

    return ddqn_rewards

# Call DDQN Training
ddqn_rewards = train_ddqn(env, num_episodes, batch_size, gamma, epsilon_start,␣
↪epsilon_end, epsilon_decay, learning_rate, target_update_period)
print("DSQN Training completed!")

```

```

Episode 1/500, Investment (actions): 209, Buying (rewards): -2.792910622432828,
Total Reward: -2.792910622432828
Target network updated at episode 0
Episode 2/500, Investment (actions): 200, Buying (rewards): -2.4703648001886904,
Total Reward: -2.4703648001886904
Episode 3/500, Investment (actions): 213, Buying (rewards): -1.9398221671581268,
Total Reward: -1.9398221671581268
Episode 4/500, Investment (actions): 201, Buying (rewards): -1.634759545326233,
Total Reward: -1.634759545326233
Episode 5/500, Investment (actions): 207, Buying (rewards): -1.101190973073244,
Total Reward: -1.101190973073244

```

Episode 6/500, Investment (actions): 175, Buying (rewards): -0.9904628287767991,
 Total Reward: -0.9904628287767991
 Episode 7/500, Investment (actions): 201, Buying (rewards): -0.7451476769056171,
 Total Reward: -0.7451476769056171
 Episode 8/500, Investment (actions): 188, Buying (rewards):
 -0.29314761725254357, Total Reward: -0.29314761725254357
 Episode 9/500, Investment (actions): 204, Buying (rewards): -0.4248990844935179,
 Total Reward: -0.4248990844935179
 Episode 10/500, Investment (actions): 194, Buying (rewards):
 -0.08071421599015594, Total Reward: -0.08071421599015594
 Episode 11/500, Investment (actions): 204, Buying (rewards):
 -0.21195507468655705, Total Reward: -0.21195507468655705
 Target network updated at episode 10
 Episode 12/500, Investment (actions): 182, Buying (rewards):
 0.07551503600552678, Total Reward: 0.07551503600552678
 Episode 13/500, Investment (actions): 205, Buying (rewards): 0.4223893079906702,
 Total Reward: 0.4223893079906702
 Episode 14/500, Investment (actions): 176, Buying (rewards): 0.7453133834060282,
 Total Reward: 0.7453133834060282
 Episode 15/500, Investment (actions): 208, Buying (rewards): 0.6706429579062387,
 Total Reward: 0.6706429579062387
 Episode 16/500, Investment (actions): 210, Buying (rewards): 0.9821087806194555,
 Total Reward: 0.9821087806194555
 Episode 17/500, Investment (actions): 182, Buying (rewards): 0.6938874679326545,
 Total Reward: 0.6938874679326545
 Episode 18/500, Investment (actions): 198, Buying (rewards): 0.4524985325115267,
 Total Reward: 0.4524985325115267
 Episode 19/500, Investment (actions): 201, Buying (rewards):
 0.002645502783707343, Total Reward: 0.002645502783707343
 Episode 20/500, Investment (actions): 193, Buying (rewards):
 0.14599169156281278, Total Reward: 0.14599169156281278
 Episode 21/500, Investment (actions): 183, Buying (rewards):
 0.01941325713414699, Total Reward: 0.01941325713414699
 Target network updated at episode 20
 Episode 22/500, Investment (actions): 209, Buying (rewards):
 0.45023694867268205, Total Reward: 0.45023694867268205
 Episode 23/500, Investment (actions): 191, Buying (rewards): 0.6158034220570698,
 Total Reward: 0.6158034220570698
 Episode 24/500, Investment (actions): 203, Buying (rewards): 0.503365378419403,
 Total Reward: 0.503365378419403
 Episode 25/500, Investment (actions): 215, Buying (rewards): 0.2269091340713203,
 Total Reward: 0.2269091340713203
 Episode 26/500, Investment (actions): 184, Buying (rewards):
 0.27839849470183253, Total Reward: 0.27839849470183253
 Episode 27/500, Investment (actions): 204, Buying (rewards):
 0.19497276097536087, Total Reward: 0.19497276097536087
 Episode 28/500, Investment (actions): 183, Buying (rewards):
 0.27726579073350877, Total Reward: 0.27726579073350877

Episode 29/500, Investment (actions): 190, Buying (rewards):
 0.03907932262518443, Total Reward: 0.03907932262518443
 Episode 30/500, Investment (actions): 212, Buying (rewards):
 -0.5026502077234909, Total Reward: -0.5026502077234909
 Episode 31/500, Investment (actions): 187, Buying (rewards):
 -0.3608929218025878, Total Reward: -0.3608929218025878
 Target network updated at episode 30
 Episode 32/500, Investment (actions): 194, Buying (rewards):
 -0.16577469877665862, Total Reward: -0.16577469877665862
 Episode 33/500, Investment (actions): 197, Buying (rewards):
 -0.5000976817682385, Total Reward: -0.5000976817682385
 Episode 34/500, Investment (actions): 183, Buying (rewards):
 -0.3801141760777682, Total Reward: -0.3801141760777682
 Episode 35/500, Investment (actions): 191, Buying (rewards):
 -0.5147895285917912, Total Reward: -0.5147895285917912
 Episode 36/500, Investment (actions): 199, Buying (rewards):
 -0.2893267328690854, Total Reward: -0.2893267328690854
 Episode 37/500, Investment (actions): 185, Buying (rewards):
 -0.2777479933865834, Total Reward: -0.2777479933865834
 Episode 38/500, Investment (actions): 224, Buying (rewards):
 -0.1970187890692614, Total Reward: -0.1970187890692614
 Episode 39/500, Investment (actions): 197, Buying (rewards):
 -0.1587632317095995, Total Reward: -0.1587632317095995
 Episode 40/500, Investment (actions): 210, Buying (rewards):
 0.03431280632503331, Total Reward: 0.03431280632503331
 Episode 41/500, Investment (actions): 197, Buying (rewards):
 0.0006354185752570629, Total Reward: 0.0006354185752570629
 Target network updated at episode 40
 Episode 42/500, Investment (actions): 209, Buying (rewards):
 -0.3372282071504742, Total Reward: -0.3372282071504742
 Episode 43/500, Investment (actions): 200, Buying (rewards):
 -0.17613574489951134, Total Reward: -0.17613574489951134
 Episode 44/500, Investment (actions): 203, Buying (rewards):
 -0.5359464202774689, Total Reward: -0.5359464202774689
 Episode 45/500, Investment (actions): 182, Buying (rewards):
 -0.7230136126745492, Total Reward: -0.7230136126745492
 Episode 46/500, Investment (actions): 204, Buying (rewards):
 -0.4136600121855736, Total Reward: -0.4136600121855736
 Episode 47/500, Investment (actions): 190, Buying (rewards):
 -0.1548608939629048, Total Reward: -0.1548608939629048
 Episode 48/500, Investment (actions): 189, Buying (rewards): 0.4025652180425823,
 Total Reward: 0.4025652180425823
 Episode 49/500, Investment (actions): 203, Buying (rewards): 0.3746971501968801,
 Total Reward: 0.3746971501968801
 Episode 50/500, Investment (actions): 187, Buying (rewards): 0.3185507708694786,
 Total Reward: 0.3185507708694786
 Episode 51/500, Investment (actions): 189, Buying (rewards): 0.4700677751097828,
 Total Reward: 0.4700677751097828

Target network updated at episode 50

Episode 52/500, Investment (actions): 206, Buying (rewards): 0.32709728088229895, Total Reward: 0.32709728088229895

Episode 53/500, Investment (actions): 213, Buying (rewards): 0.24920884691528045, Total Reward: 0.24920884691528045

Episode 54/500, Investment (actions): 206, Buying (rewards): 0.0822480942151742, Total Reward: 0.0822480942151742

Episode 55/500, Investment (actions): 209, Buying (rewards): 0.36971730763616506, Total Reward: 0.36971730763616506

Episode 56/500, Investment (actions): 218, Buying (rewards): 0.4868232093285769, Total Reward: 0.4868232093285769

Episode 57/500, Investment (actions): 179, Buying (rewards): 0.2965216306038201, Total Reward: 0.2965216306038201

Episode 58/500, Investment (actions): 192, Buying (rewards): 0.41775021981447935, Total Reward: 0.41775021981447935

Episode 59/500, Investment (actions): 195, Buying (rewards): 0.373357457043312, Total Reward: 0.373357457043312

Episode 60/500, Investment (actions): 186, Buying (rewards): 0.2881210398336407, Total Reward: 0.2881210398336407

Episode 61/500, Investment (actions): 209, Buying (rewards): 0.4123677609022707, Total Reward: 0.4123677609022707

Target network updated at episode 60

Episode 62/500, Investment (actions): 215, Buying (rewards): 0.0997148520546034, Total Reward: 0.0997148520546034

Episode 63/500, Investment (actions): 197, Buying (rewards): 0.27698601443262305, Total Reward: 0.27698601443262305

Episode 64/500, Investment (actions): 199, Buying (rewards): 0.2877799230918754, Total Reward: 0.2877799230918754

Episode 65/500, Investment (actions): 200, Buying (rewards): 0.22538512709434144, Total Reward: 0.22538512709434144

Episode 66/500, Investment (actions): 192, Buying (rewards): 0.3576071842107922, Total Reward: 0.3576071842107922

Episode 67/500, Investment (actions): 206, Buying (rewards): 0.36041779059451073, Total Reward: 0.36041779059451073

Episode 68/500, Investment (actions): 190, Buying (rewards): 0.6946744651068002, Total Reward: 0.6946744651068002

Episode 69/500, Investment (actions): 216, Buying (rewards): 0.4342666584998369, Total Reward: 0.4342666584998369

Episode 70/500, Investment (actions): 204, Buying (rewards): 0.7282604635111056, Total Reward: 0.7282604635111056

Episode 71/500, Investment (actions): 216, Buying (rewards): 0.5742183758411556, Total Reward: 0.5742183758411556

Target network updated at episode 70

Episode 72/500, Investment (actions): 184, Buying (rewards): 0.25401000306010246, Total Reward: 0.25401000306010246

Episode 73/500, Investment (actions): 213, Buying (rewards): 0.0320034094620496, Total Reward: 0.0320034094620496

Episode 74/500, Investment (actions): 198, Buying (rewards):

0.19475985586177558, Total Reward: 0.19475985586177558
 Episode 75/500, Investment (actions): 210, Buying (rewards):
 0.27134946384467185, Total Reward: 0.27134946384467185
 Episode 76/500, Investment (actions): 181, Buying (rewards):
 0.43013684591278434, Total Reward: 0.43013684591278434
 Episode 77/500, Investment (actions): 190, Buying (rewards): 0.3564418666064739,
 Total Reward: 0.3564418666064739
 Episode 78/500, Investment (actions): 210, Buying (rewards):
 -0.06111058231908828, Total Reward: -0.06111058231908828
 Episode 79/500, Investment (actions): 198, Buying (rewards):
 -0.11843346757814288, Total Reward: -0.11843346757814288
 Episode 80/500, Investment (actions): 210, Buying (rewards):
 -0.08083893707953393, Total Reward: -0.08083893707953393
 Episode 81/500, Investment (actions): 209, Buying (rewards): -0.048801374505274,
 Total Reward: -0.048801374505274
 Target network updated at episode 80
 Episode 82/500, Investment (actions): 204, Buying (rewards):
 -0.09973032693960704, Total Reward: -0.09973032693960704
 Episode 83/500, Investment (actions): 204, Buying (rewards):
 0.14518484263680875, Total Reward: 0.14518484263680875
 Episode 84/500, Investment (actions): 192, Buying (rewards):
 0.17878630640916526, Total Reward: 0.17878630640916526
 Episode 85/500, Investment (actions): 204, Buying (rewards):
 0.38549018767662346, Total Reward: 0.38549018767662346
 Episode 86/500, Investment (actions): 214, Buying (rewards): 0.5793042881414294,
 Total Reward: 0.5793042881414294
 Episode 87/500, Investment (actions): 203, Buying (rewards): 0.7372207073494792,
 Total Reward: 0.7372207073494792
 Episode 88/500, Investment (actions): 200, Buying (rewards):
 0.18223285186104476, Total Reward: 0.18223285186104476
 Episode 89/500, Investment (actions): 195, Buying (rewards):
 -0.05129191279411316, Total Reward: -0.05129191279411316
 Episode 90/500, Investment (actions): 195, Buying (rewards):
 0.21853236202150583, Total Reward: 0.21853236202150583
 Episode 91/500, Investment (actions): 201, Buying (rewards):
 0.12220972031354904, Total Reward: 0.12220972031354904
 Target network updated at episode 90
 Episode 92/500, Investment (actions): 212, Buying (rewards):
 -0.14259965904057026, Total Reward: -0.14259965904057026
 Episode 93/500, Investment (actions): 210, Buying (rewards):
 0.09049658849835396, Total Reward: 0.09049658849835396
 Episode 94/500, Investment (actions): 189, Buying (rewards):
 0.06522944569587708, Total Reward: 0.06522944569587708
 Episode 95/500, Investment (actions): 192, Buying (rewards):
 -0.01391345341107808, Total Reward: -0.01391345341107808
 Episode 96/500, Investment (actions): 212, Buying (rewards): 0.137170977628557,
 Total Reward: 0.137170977628557
 Episode 97/500, Investment (actions): 214, Buying (rewards):

-0.0017583632288733497, Total Reward: -0.0017583632288733497
 Episode 98/500, Investment (actions): 197, Buying (rewards):
 -0.017082132602808997, Total Reward: -0.017082132602808997
 Episode 99/500, Investment (actions): 204, Buying (rewards): 0.0934654549928382,
 Total Reward: 0.0934654549928382
 Episode 100/500, Investment (actions): 195, Buying (rewards):
 0.3416121157351881, Total Reward: 0.3416121157351881
 Episode 101/500, Investment (actions): 195, Buying (rewards):
 0.3822057214565575, Total Reward: 0.3822057214565575
 Target network updated at episode 100
 Episode 102/500, Investment (actions): 206, Buying (rewards):
 0.2594216151628643, Total Reward: 0.2594216151628643
 Episode 103/500, Investment (actions): 193, Buying (rewards):
 0.19987928122282028, Total Reward: 0.19987928122282028
 Episode 104/500, Investment (actions): 202, Buying (rewards):
 0.22012130683287978, Total Reward: 0.22012130683287978
 Episode 105/500, Investment (actions): 207, Buying (rewards):
 0.36735455063171685, Total Reward: 0.36735455063171685
 Episode 106/500, Investment (actions): 178, Buying (rewards): 0.323957514250651,
 Total Reward: 0.323957514250651
 Episode 107/500, Investment (actions): 201, Buying (rewards):
 0.3201862066052854, Total Reward: 0.3201862066052854
 Episode 108/500, Investment (actions): 197, Buying (rewards):
 0.3911686409264803, Total Reward: 0.3911686409264803
 Episode 109/500, Investment (actions): 209, Buying (rewards):
 0.20926478318870068, Total Reward: 0.20926478318870068
 Episode 110/500, Investment (actions): 214, Buying (rewards):
 0.25646232813596725, Total Reward: 0.25646232813596725
 Episode 111/500, Investment (actions): 194, Buying (rewards):
 0.2261010219808668, Total Reward: 0.2261010219808668
 Target network updated at episode 110
 Episode 112/500, Investment (actions): 194, Buying (rewards):
 0.17671465035527945, Total Reward: 0.17671465035527945
 Episode 113/500, Investment (actions): 205, Buying (rewards):
 0.24910758645273745, Total Reward: 0.24910758645273745
 Episode 114/500, Investment (actions): 217, Buying (rewards):
 0.023897249717265368, Total Reward: 0.023897249717265368
 Episode 115/500, Investment (actions): 202, Buying (rewards):
 -0.027440521866083145, Total Reward: -0.027440521866083145
 Episode 116/500, Investment (actions): 211, Buying (rewards):
 -0.09341023862361908, Total Reward: -0.09341023862361908
 Episode 117/500, Investment (actions): 197, Buying (rewards):
 -0.05741872638463974, Total Reward: -0.05741872638463974
 Episode 118/500, Investment (actions): 207, Buying (rewards):
 -0.4891727091744542, Total Reward: -0.4891727091744542
 Episode 119/500, Investment (actions): 205, Buying (rewards):
 -0.22196234110742807, Total Reward: -0.22196234110742807
 Episode 120/500, Investment (actions): 205, Buying (rewards):

-0.2650926047936082, Total Reward: -0.2650926047936082
 Episode 121/500, Investment (actions): 193, Buying (rewards):
 -0.12245113169774413, Total Reward: -0.12245113169774413
 Target network updated at episode 120
 Episode 122/500, Investment (actions): 190, Buying (rewards):
 0.17547083995305002, Total Reward: 0.17547083995305002
 Episode 123/500, Investment (actions): 189, Buying (rewards):
 0.011858259560540318, Total Reward: 0.011858259560540318
 Episode 124/500, Investment (actions): 205, Buying (rewards):
 0.14885538260568865, Total Reward: 0.14885538260568865
 Episode 125/500, Investment (actions): 211, Buying (rewards):
 0.03996117445058189, Total Reward: 0.03996117445058189
 Episode 126/500, Investment (actions): 213, Buying (rewards):
 0.04093020735308528, Total Reward: 0.04093020735308528
 Episode 127/500, Investment (actions): 189, Buying (rewards):
 -0.10795641131699085, Total Reward: -0.10795641131699085
 Episode 128/500, Investment (actions): 200, Buying (rewards):
 -0.20615029938926455, Total Reward: -0.20615029938926455
 Episode 129/500, Investment (actions): 205, Buying (rewards):
 -0.1926707021339098, Total Reward: -0.1926707021339098
 Episode 130/500, Investment (actions): 219, Buying (rewards):
 -0.045188436950411415, Total Reward: -0.045188436950411415
 Episode 131/500, Investment (actions): 180, Buying (rewards):
 -0.21861680503934622, Total Reward: -0.21861680503934622
 Target network updated at episode 130
 Episode 132/500, Investment (actions): 200, Buying (rewards):
 -0.1666586622595787, Total Reward: -0.1666586622595787
 Episode 133/500, Investment (actions): 197, Buying (rewards):
 -0.30823755636811256, Total Reward: -0.30823755636811256
 Episode 134/500, Investment (actions): 202, Buying (rewards):
 -0.32402512105181813, Total Reward: -0.32402512105181813
 Episode 135/500, Investment (actions): 191, Buying (rewards):
 -0.2365263751707971, Total Reward: -0.2365263751707971
 Episode 136/500, Investment (actions): 207, Buying (rewards):
 -0.23139667324721813, Total Reward: -0.23139667324721813
 Episode 137/500, Investment (actions): 214, Buying (rewards):
 -0.40213779665646143, Total Reward: -0.40213779665646143
 Episode 138/500, Investment (actions): 192, Buying (rewards):
 -0.4818592071533203, Total Reward: -0.4818592071533203
 Episode 139/500, Investment (actions): 192, Buying (rewards):
 -0.09659599512815475, Total Reward: -0.09659599512815475
 Episode 140/500, Investment (actions): 198, Buying (rewards):
 -0.14883194025605917, Total Reward: -0.14883194025605917
 Episode 141/500, Investment (actions): 210, Buying (rewards):
 -0.07023491151630878, Total Reward: -0.07023491151630878
 Target network updated at episode 140
 Episode 142/500, Investment (actions): 198, Buying (rewards):
 0.09080336429178715, Total Reward: 0.09080336429178715

Episode 143/500, Investment (actions): 207, Buying (rewards):
 0.11673148348927498, Total Reward: 0.11673148348927498
 Episode 144/500, Investment (actions): 232, Buying (rewards):
 -0.29053943417966366, Total Reward: -0.29053943417966366
 Episode 145/500, Investment (actions): 200, Buying (rewards):
 -0.1561543890275061, Total Reward: -0.1561543890275061
 Episode 146/500, Investment (actions): 198, Buying (rewards):
 0.09496734326239675, Total Reward: 0.09496734326239675
 Episode 147/500, Investment (actions): 200, Buying (rewards):
 0.05877534265164286, Total Reward: 0.05877534265164286
 Episode 148/500, Investment (actions): 212, Buying (rewards):
 -0.19591061165556312, Total Reward: -0.19591061165556312
 Episode 149/500, Investment (actions): 210, Buying (rewards):
 -0.05162263195961714, Total Reward: -0.05162263195961714
 Episode 150/500, Investment (actions): 200, Buying (rewards):
 -0.13436082331463695, Total Reward: -0.13436082331463695
 Episode 151/500, Investment (actions): 190, Buying (rewards):
 0.0016360897570848465, Total Reward: 0.0016360897570848465
 Target network updated at episode 150
 Episode 152/500, Investment (actions): 199, Buying (rewards):
 -0.15112202800810337, Total Reward: -0.15112202800810337
 Episode 153/500, Investment (actions): 198, Buying (rewards):
 -0.013273786869831383, Total Reward: -0.013273786869831383
 Episode 154/500, Investment (actions): 199, Buying (rewards):
 -0.1517979452619329, Total Reward: -0.1517979452619329
 Episode 155/500, Investment (actions): 206, Buying (rewards):
 -0.19622092152712867, Total Reward: -0.19622092152712867
 Episode 156/500, Investment (actions): 208, Buying (rewards):
 -0.06582440971396863, Total Reward: -0.06582440971396863
 Episode 157/500, Investment (actions): 200, Buying (rewards):
 -0.005117885186336935, Total Reward: -0.005117885186336935
 Episode 158/500, Investment (actions): 184, Buying (rewards):
 -0.05233720917021856, Total Reward: -0.05233720917021856
 Episode 159/500, Investment (actions): 196, Buying (rewards):
 0.21805572602897882, Total Reward: 0.21805572602897882
 Episode 160/500, Investment (actions): 209, Buying (rewards):
 0.14926026295870543, Total Reward: 0.14926026295870543
 Episode 161/500, Investment (actions): 201, Buying (rewards):
 0.00552508607506752, Total Reward: 0.00552508607506752
 Target network updated at episode 160
 Episode 162/500, Investment (actions): 203, Buying (rewards):
 0.07303991541266441, Total Reward: 0.07303991541266441
 Episode 163/500, Investment (actions): 215, Buying (rewards):
 0.1479331876616925, Total Reward: 0.1479331876616925
 Episode 164/500, Investment (actions): 197, Buying (rewards):
 0.04204239649698138, Total Reward: 0.04204239649698138
 Episode 165/500, Investment (actions): 200, Buying (rewards):
 0.1536361388862133, Total Reward: 0.1536361388862133

Episode 166/500, Investment (actions): 193, Buying (rewards):
 0.025817466201260686, Total Reward: 0.025817466201260686
 Episode 167/500, Investment (actions): 209, Buying (rewards):
 -0.006869426928460598, Total Reward: -0.006869426928460598
 Episode 168/500, Investment (actions): 210, Buying (rewards):
 -0.20987785141915083, Total Reward: -0.20987785141915083
 Episode 169/500, Investment (actions): 198, Buying (rewards):
 0.14848236343823373, Total Reward: 0.14848236343823373
 Episode 170/500, Investment (actions): 216, Buying (rewards):
 0.19494226295500994, Total Reward: 0.19494226295500994
 Episode 171/500, Investment (actions): 206, Buying (rewards):
 0.06884249998256564, Total Reward: 0.06884249998256564
 Target network updated at episode 170
 Episode 172/500, Investment (actions): 217, Buying (rewards):
 -0.15963758737780154, Total Reward: -0.15963758737780154
 Episode 173/500, Investment (actions): 202, Buying (rewards):
 -0.3113933112472296, Total Reward: -0.3113933112472296
 Episode 174/500, Investment (actions): 187, Buying (rewards):
 -0.05120178498327732, Total Reward: -0.05120178498327732
 Episode 175/500, Investment (actions): 198, Buying (rewards):
 -0.1911706207320094, Total Reward: -0.1911706207320094
 Episode 176/500, Investment (actions): 198, Buying (rewards):
 -0.21969629311934114, Total Reward: -0.21969629311934114
 Episode 177/500, Investment (actions): 196, Buying (rewards):
 -0.09371779533103108, Total Reward: -0.09371779533103108
 Episode 178/500, Investment (actions): 198, Buying (rewards):
 0.010538440197706223, Total Reward: 0.010538440197706223
 Episode 179/500, Investment (actions): 189, Buying (rewards):
 0.10857804864645004, Total Reward: 0.10857804864645004
 Episode 180/500, Investment (actions): 197, Buying (rewards):
 0.0820580581203103, Total Reward: 0.0820580581203103
 Episode 181/500, Investment (actions): 202, Buying (rewards):
 -0.1263826466165483, Total Reward: -0.1263826466165483
 Target network updated at episode 180
 Episode 182/500, Investment (actions): 197, Buying (rewards):
 -0.0016977260820567608, Total Reward: -0.0016977260820567608
 Episode 183/500, Investment (actions): 191, Buying (rewards):
 -0.05593139724805951, Total Reward: -0.05593139724805951
 Episode 184/500, Investment (actions): 199, Buying (rewards):
 0.006149019580334425, Total Reward: 0.006149019580334425
 Episode 185/500, Investment (actions): 206, Buying (rewards):
 -0.1065343115478754, Total Reward: -0.1065343115478754
 Episode 186/500, Investment (actions): 201, Buying (rewards):
 -0.0802013617940247, Total Reward: -0.0802013617940247
 Episode 187/500, Investment (actions): 195, Buying (rewards):
 -0.01049633976072073, Total Reward: -0.01049633976072073
 Episode 188/500, Investment (actions): 202, Buying (rewards):
 0.0417454494163394, Total Reward: 0.0417454494163394

Episode 189/500, Investment (actions): 190, Buying (rewards):
 0.024720282410271466, Total Reward: 0.024720282410271466
 Episode 190/500, Investment (actions): 199, Buying (rewards):
 0.3397246031090617, Total Reward: 0.3397246031090617
 Episode 191/500, Investment (actions): 193, Buying (rewards):
 0.43226582487113774, Total Reward: 0.43226582487113774
 Target network updated at episode 190
 Episode 192/500, Investment (actions): 195, Buying (rewards):
 0.21468255622312427, Total Reward: 0.21468255622312427
 Episode 193/500, Investment (actions): 205, Buying (rewards):
 0.06428868044167757, Total Reward: 0.06428868044167757
 Episode 194/500, Investment (actions): 197, Buying (rewards):
 0.043051589746028185, Total Reward: 0.043051589746028185
 Episode 195/500, Investment (actions): 208, Buying (rewards):
 0.0009183380752801895, Total Reward: 0.0009183380752801895
 Episode 196/500, Investment (actions): 209, Buying (rewards):
 0.05289152357727289, Total Reward: 0.05289152357727289
 Episode 197/500, Investment (actions): 197, Buying (rewards):
 0.16092591593042016, Total Reward: 0.16092591593042016
 Episode 198/500, Investment (actions): 185, Buying (rewards):
 0.021197878289967775, Total Reward: 0.021197878289967775
 Episode 199/500, Investment (actions): 210, Buying (rewards):
 0.051414423156529665, Total Reward: 0.051414423156529665
 Episode 200/500, Investment (actions): 195, Buying (rewards):
 -0.2535686627961695, Total Reward: -0.2535686627961695
 Episode 201/500, Investment (actions): 200, Buying (rewards):
 -0.1661314619705081, Total Reward: -0.1661314619705081
 Target network updated at episode 200
 Episode 202/500, Investment (actions): 213, Buying (rewards):
 -0.543455564416945, Total Reward: -0.543455564416945
 Episode 203/500, Investment (actions): 208, Buying (rewards):
 -0.5120081440545619, Total Reward: -0.5120081440545619
 Episode 204/500, Investment (actions): 199, Buying (rewards):
 -0.4252423176076263, Total Reward: -0.4252423176076263
 Episode 205/500, Investment (actions): 189, Buying (rewards):
 -0.2601454542018473, Total Reward: -0.2601454542018473
 Episode 206/500, Investment (actions): 197, Buying (rewards):
 -0.12684918893501163, Total Reward: -0.12684918893501163
 Episode 207/500, Investment (actions): 189, Buying (rewards):
 -0.019589459290727973, Total Reward: -0.019589459290727973
 Episode 208/500, Investment (actions): 191, Buying (rewards):
 0.05461192550137639, Total Reward: 0.05461192550137639
 Episode 209/500, Investment (actions): 204, Buying (rewards):
 -0.09512638067826629, Total Reward: -0.09512638067826629
 Episode 210/500, Investment (actions): 198, Buying (rewards):
 -0.1377545716241002, Total Reward: -0.1377545716241002
 Episode 211/500, Investment (actions): 200, Buying (rewards):
 0.03751545026898384, Total Reward: 0.03751545026898384

Target network updated at episode 210

Episode 212/500, Investment (actions): 197, Buying (rewards): 0.02088003233075142, Total Reward: 0.02088003233075142

Episode 213/500, Investment (actions): 202, Buying (rewards): -0.08745585288852453, Total Reward: -0.08745585288852453

Episode 214/500, Investment (actions): 196, Buying (rewards): 0.005708350799977779, Total Reward: 0.005708350799977779

Episode 215/500, Investment (actions): 205, Buying (rewards): -0.0928532900288701, Total Reward: -0.0928532900288701

Episode 216/500, Investment (actions): 191, Buying (rewards): 0.10747059620916843, Total Reward: 0.10747059620916843

Episode 217/500, Investment (actions): 207, Buying (rewards): 0.15749601484276354, Total Reward: 0.15749601484276354

Episode 218/500, Investment (actions): 200, Buying (rewards): 0.22915948391892016, Total Reward: 0.22915948391892016

Episode 219/500, Investment (actions): 205, Buying (rewards): 0.1692647624295205, Total Reward: 0.1692647624295205

Episode 220/500, Investment (actions): 194, Buying (rewards): 0.20203682617284358, Total Reward: 0.20203682617284358

Episode 221/500, Investment (actions): 209, Buying (rewards): 0.17262084782123566, Total Reward: 0.17262084782123566

Target network updated at episode 220

Episode 222/500, Investment (actions): 210, Buying (rewards): 0.09418034111149609, Total Reward: 0.09418034111149609

Episode 223/500, Investment (actions): 196, Buying (rewards): 0.42906804708763957, Total Reward: 0.42906804708763957

Episode 224/500, Investment (actions): 205, Buying (rewards): 0.21504656318575144, Total Reward: 0.21504656318575144

Episode 225/500, Investment (actions): 200, Buying (rewards): 0.24502793047577143, Total Reward: 0.24502793047577143

Episode 226/500, Investment (actions): 196, Buying (rewards): 0.2248600427992642, Total Reward: 0.2248600427992642

Episode 227/500, Investment (actions): 219, Buying (rewards): 0.10157467576209456, Total Reward: 0.10157467576209456

Episode 228/500, Investment (actions): 190, Buying (rewards): 0.260125178610906, Total Reward: 0.260125178610906

Episode 229/500, Investment (actions): 204, Buying (rewards): 0.05006471648812294, Total Reward: 0.05006471648812294

Episode 230/500, Investment (actions): 202, Buying (rewards): -0.004983936785720289, Total Reward: -0.004983936785720289

Episode 231/500, Investment (actions): 207, Buying (rewards): 0.05060516856610775, Total Reward: 0.05060516856610775

Target network updated at episode 230

Episode 232/500, Investment (actions): 210, Buying (rewards): -0.05363688536453992, Total Reward: -0.05363688536453992

Episode 233/500, Investment (actions): 202, Buying (rewards): -0.04765979980584234, Total Reward: -0.04765979980584234

Episode 234/500, Investment (actions): 197, Buying (rewards):

0.03524680552072823, Total Reward: 0.03524680552072823
 Episode 235/500, Investment (actions): 204, Buying (rewards):
 0.04951438249554485, Total Reward: 0.04951438249554485
 Episode 236/500, Investment (actions): 203, Buying (rewards):
 0.28909923578612506, Total Reward: 0.28909923578612506
 Episode 237/500, Investment (actions): 202, Buying (rewards):
 0.2576039284467697, Total Reward: 0.2576039284467697
 Episode 238/500, Investment (actions): 206, Buying (rewards):
 0.27499415073543787, Total Reward: 0.27499415073543787
 Episode 239/500, Investment (actions): 205, Buying (rewards):
 0.21934672351926565, Total Reward: 0.21934672351926565
 Episode 240/500, Investment (actions): 208, Buying (rewards):
 0.13612590730190277, Total Reward: 0.13612590730190277
 Episode 241/500, Investment (actions): 197, Buying (rewards):
 0.1117380007635802, Total Reward: 0.1117380007635802
 Target network updated at episode 240
 Episode 242/500, Investment (actions): 205, Buying (rewards):
 0.15866222488693893, Total Reward: 0.15866222488693893
 Episode 243/500, Investment (actions): 195, Buying (rewards):
 0.08959746942855418, Total Reward: 0.08959746942855418
 Episode 244/500, Investment (actions): 202, Buying (rewards):
 0.09151457459665835, Total Reward: 0.09151457459665835
 Episode 245/500, Investment (actions): 197, Buying (rewards):
 -0.09818936162628233, Total Reward: -0.09818936162628233
 Episode 246/500, Investment (actions): 207, Buying (rewards):
 -0.015344669809564948, Total Reward: -0.015344669809564948
 Episode 247/500, Investment (actions): 203, Buying (rewards):
 -0.05638419999741018, Total Reward: -0.05638419999741018
 Episode 248/500, Investment (actions): 208, Buying (rewards):
 -0.06913450756110251, Total Reward: -0.06913450756110251
 Episode 249/500, Investment (actions): 209, Buying (rewards):
 -0.15173598751425743, Total Reward: -0.15173598751425743
 Episode 250/500, Investment (actions): 197, Buying (rewards):
 -0.24923413037322462, Total Reward: -0.24923413037322462
 Episode 251/500, Investment (actions): 192, Buying (rewards):
 -0.45805950928479433, Total Reward: -0.45805950928479433
 Target network updated at episode 250
 Episode 252/500, Investment (actions): 192, Buying (rewards):
 -0.5251943003386259, Total Reward: -0.5251943003386259
 Episode 253/500, Investment (actions): 195, Buying (rewards):
 -0.3324667327105999, Total Reward: -0.3324667327105999
 Episode 254/500, Investment (actions): 200, Buying (rewards):
 -0.3171618948690593, Total Reward: -0.3171618948690593
 Episode 255/500, Investment (actions): 200, Buying (rewards):
 -0.4290684061124921, Total Reward: -0.4290684061124921
 Episode 256/500, Investment (actions): 207, Buying (rewards):
 -0.5339040388353169, Total Reward: -0.5339040388353169
 Episode 257/500, Investment (actions): 196, Buying (rewards):

-0.5560530270449817, Total Reward: -0.5560530270449817
 Episode 258/500, Investment (actions): 198, Buying (rewards):
 -0.41085356986150146, Total Reward: -0.41085356986150146
 Episode 259/500, Investment (actions): 206, Buying (rewards):
 -0.3477558103040792, Total Reward: -0.3477558103040792
 Episode 260/500, Investment (actions): 200, Buying (rewards):
 -0.530479052249575, Total Reward: -0.530479052249575
 Episode 261/500, Investment (actions): 190, Buying (rewards):
 -0.3459898363798857, Total Reward: -0.3459898363798857
 Target network updated at episode 260
 Episode 262/500, Investment (actions): 201, Buying (rewards):
 -0.2634064191952348, Total Reward: -0.2634064191952348
 Episode 263/500, Investment (actions): 200, Buying (rewards):
 -0.264953076839447, Total Reward: -0.264953076839447
 Episode 264/500, Investment (actions): 206, Buying (rewards):
 -0.12645118683576584, Total Reward: -0.12645118683576584
 Episode 265/500, Investment (actions): 208, Buying (rewards):
 -0.21299030631780624, Total Reward: -0.21299030631780624
 Episode 266/500, Investment (actions): 198, Buying (rewards):
 -0.1934571024030447, Total Reward: -0.1934571024030447
 Episode 267/500, Investment (actions): 204, Buying (rewards):
 -0.10493071842938662, Total Reward: -0.10493071842938662
 Episode 268/500, Investment (actions): 196, Buying (rewards):
 0.14518013782799244, Total Reward: 0.14518013782799244
 Episode 269/500, Investment (actions): 199, Buying (rewards):
 0.3021236741915345, Total Reward: 0.3021236741915345
 Episode 270/500, Investment (actions): 193, Buying (rewards):
 0.23471283726394176, Total Reward: 0.23471283726394176
 Episode 271/500, Investment (actions): 206, Buying (rewards):
 0.24852752313017845, Total Reward: 0.24852752313017845
 Target network updated at episode 270
 Episode 272/500, Investment (actions): 182, Buying (rewards):
 0.30856269784271717, Total Reward: 0.30856269784271717
 Episode 273/500, Investment (actions): 200, Buying (rewards):
 0.3051501838490367, Total Reward: 0.3051501838490367
 Episode 274/500, Investment (actions): 198, Buying (rewards):
 0.1746355565264821, Total Reward: 0.1746355565264821
 Episode 275/500, Investment (actions): 200, Buying (rewards):
 0.0396235475782305, Total Reward: 0.0396235475782305
 Episode 276/500, Investment (actions): 200, Buying (rewards):
 -0.001996713108383119, Total Reward: -0.001996713108383119
 Episode 277/500, Investment (actions): 197, Buying (rewards):
 -0.2116530837956816, Total Reward: -0.2116530837956816
 Episode 278/500, Investment (actions): 196, Buying (rewards):
 -0.36861759377643466, Total Reward: -0.36861759377643466
 Episode 279/500, Investment (actions): 197, Buying (rewards):
 -0.47186648030765355, Total Reward: -0.47186648030765355
 Episode 280/500, Investment (actions): 201, Buying (rewards):

-0.43295127409510314, Total Reward: -0.43295127409510314
 Episode 281/500, Investment (actions): 200, Buying (rewards):
 -0.25328189227730036, Total Reward: -0.25328189227730036
 Target network updated at episode 280
 Episode 282/500, Investment (actions): 206, Buying (rewards):
 -0.4254431212320924, Total Reward: -0.4254431212320924
 Episode 283/500, Investment (actions): 190, Buying (rewards):
 -0.3762362999841571, Total Reward: -0.3762362999841571
 Episode 284/500, Investment (actions): 201, Buying (rewards):
 -0.38800061866641045, Total Reward: -0.38800061866641045
 Episode 285/500, Investment (actions): 196, Buying (rewards):
 -0.2647106982767582, Total Reward: -0.2647106982767582
 Episode 286/500, Investment (actions): 201, Buying (rewards):
 -0.45805505383759737, Total Reward: -0.45805505383759737
 Episode 287/500, Investment (actions): 199, Buying (rewards):
 -0.4165704483166337, Total Reward: -0.4165704483166337
 Episode 288/500, Investment (actions): 201, Buying (rewards):
 -0.42644433211535215, Total Reward: -0.42644433211535215
 Episode 289/500, Investment (actions): 194, Buying (rewards):
 -0.19252635026350617, Total Reward: -0.19252635026350617
 Episode 290/500, Investment (actions): 207, Buying (rewards):
 -0.23875372484326363, Total Reward: -0.23875372484326363
 Episode 291/500, Investment (actions): 206, Buying (rewards):
 -0.2766293724998832, Total Reward: -0.2766293724998832
 Target network updated at episode 290
 Episode 292/500, Investment (actions): 195, Buying (rewards):
 -0.18166069174185395, Total Reward: -0.18166069174185395
 Episode 293/500, Investment (actions): 209, Buying (rewards):
 -0.05039865477010608, Total Reward: -0.05039865477010608
 Episode 294/500, Investment (actions): 205, Buying (rewards):
 0.10726168192923069, Total Reward: 0.10726168192923069
 Episode 295/500, Investment (actions): 204, Buying (rewards):
 0.24150577653199434, Total Reward: 0.24150577653199434
 Episode 296/500, Investment (actions): 195, Buying (rewards):
 0.3320693392306566, Total Reward: 0.3320693392306566
 Episode 297/500, Investment (actions): 201, Buying (rewards):
 0.1814979426562786, Total Reward: 0.1814979426562786
 Episode 298/500, Investment (actions): 200, Buying (rewards):
 0.22627671994268894, Total Reward: 0.22627671994268894
 Episode 299/500, Investment (actions): 198, Buying (rewards):
 0.15225785318762064, Total Reward: 0.15225785318762064
 Episode 300/500, Investment (actions): 191, Buying (rewards):
 0.25257807364687324, Total Reward: 0.25257807364687324
 Episode 301/500, Investment (actions): 200, Buying (rewards):
 0.01604568213224411, Total Reward: 0.01604568213224411
 Target network updated at episode 300
 Episode 302/500, Investment (actions): 198, Buying (rewards):
 0.24948999099433422, Total Reward: 0.24948999099433422

Episode 303/500, Investment (actions): 195, Buying (rewards):
 0.10502928867936134, Total Reward: 0.10502928867936134
 Episode 304/500, Investment (actions): 204, Buying (rewards):
 0.0075984615832567215, Total Reward: 0.0075984615832567215
 Episode 305/500, Investment (actions): 193, Buying (rewards):
 0.1313840551301837, Total Reward: 0.1313840551301837
 Episode 306/500, Investment (actions): 208, Buying (rewards):
 -0.03296773415058851, Total Reward: -0.03296773415058851
 Episode 307/500, Investment (actions): 200, Buying (rewards):
 -0.29674133844673634, Total Reward: -0.29674133844673634
 Episode 308/500, Investment (actions): 203, Buying (rewards):
 -0.29010515846312046, Total Reward: -0.29010515846312046
 Episode 309/500, Investment (actions): 208, Buying (rewards):
 -0.21502468781545758, Total Reward: -0.21502468781545758
 Episode 310/500, Investment (actions): 197, Buying (rewards):
 -0.1826087930239737, Total Reward: -0.1826087930239737
 Episode 311/500, Investment (actions): 193, Buying (rewards):
 -0.08605130854994059, Total Reward: -0.08605130854994059
 Target network updated at episode 310
 Episode 312/500, Investment (actions): 208, Buying (rewards):
 -0.22694043023511767, Total Reward: -0.22694043023511767
 Episode 313/500, Investment (actions): 198, Buying (rewards):
 -0.3789180633611977, Total Reward: -0.3789180633611977
 Episode 314/500, Investment (actions): 203, Buying (rewards):
 -0.21118659060448408, Total Reward: -0.21118659060448408
 Episode 315/500, Investment (actions): 207, Buying (rewards):
 -0.26044839434325695, Total Reward: -0.26044839434325695
 Episode 316/500, Investment (actions): 197, Buying (rewards):
 -0.376256212242879, Total Reward: -0.376256212242879
 Episode 317/500, Investment (actions): 204, Buying (rewards):
 -0.34367930609732866, Total Reward: -0.34367930609732866
 Episode 318/500, Investment (actions): 204, Buying (rewards):
 -0.264441329985857, Total Reward: -0.264441329985857
 Episode 319/500, Investment (actions): 193, Buying (rewards):
 -0.3063700683414936, Total Reward: -0.3063700683414936
 Episode 320/500, Investment (actions): 207, Buying (rewards):
 -0.17821779544465244, Total Reward: -0.17821779544465244
 Episode 321/500, Investment (actions): 207, Buying (rewards):
 -0.18371787737123668, Total Reward: -0.18371787737123668
 Target network updated at episode 320
 Episode 322/500, Investment (actions): 196, Buying (rewards):
 -0.20218689180910587, Total Reward: -0.20218689180910587
 Episode 323/500, Investment (actions): 210, Buying (rewards):
 -0.3536609970033169, Total Reward: -0.3536609970033169
 Episode 324/500, Investment (actions): 192, Buying (rewards):
 -0.1979108860832639, Total Reward: -0.1979108860832639
 Episode 325/500, Investment (actions): 202, Buying (rewards):
 -0.08506045996909961, Total Reward: -0.08506045996909961

Episode 326/500, Investment (actions): 208, Buying (rewards):
 -0.04663357883691788, Total Reward: -0.04663357883691788
 Episode 327/500, Investment (actions): 196, Buying (rewards):
 0.03655050881206989, Total Reward: 0.03655050881206989
 Episode 328/500, Investment (actions): 200, Buying (rewards):
 0.14442489482462406, Total Reward: 0.14442489482462406
 Episode 329/500, Investment (actions): 205, Buying (rewards):
 0.10012434236705303, Total Reward: 0.10012434236705303
 Episode 330/500, Investment (actions): 205, Buying (rewards):
 0.10331492323894054, Total Reward: 0.10331492323894054
 Episode 331/500, Investment (actions): 195, Buying (rewards):
 0.1518260476877913, Total Reward: 0.1518260476877913
 Target network updated at episode 330
 Episode 332/500, Investment (actions): 203, Buying (rewards):
 0.09396156168077141, Total Reward: 0.09396156168077141
 Episode 333/500, Investment (actions): 204, Buying (rewards):
 0.23469683155417442, Total Reward: 0.23469683155417442
 Episode 334/500, Investment (actions): 196, Buying (rewards):
 0.1537655908614397, Total Reward: 0.1537655908614397
 Episode 335/500, Investment (actions): 197, Buying (rewards):
 0.03339477814733982, Total Reward: 0.03339477814733982
 Episode 336/500, Investment (actions): 207, Buying (rewards):
 -0.020090796053409576, Total Reward: -0.020090796053409576
 Episode 337/500, Investment (actions): 208, Buying (rewards):
 -0.05364898033440113, Total Reward: -0.05364898033440113
 Episode 338/500, Investment (actions): 200, Buying (rewards):
 -0.11391244805417955, Total Reward: -0.11391244805417955
 Episode 339/500, Investment (actions): 214, Buying (rewards):
 -0.09909304231405258, Total Reward: -0.09909304231405258
 Episode 340/500, Investment (actions): 193, Buying (rewards):
 -0.03639437444508076, Total Reward: -0.03639437444508076
 Episode 341/500, Investment (actions): 204, Buying (rewards):
 -0.06678391061723232, Total Reward: -0.06678391061723232
 Target network updated at episode 340
 Episode 342/500, Investment (actions): 194, Buying (rewards):
 -0.08527384884655476, Total Reward: -0.08527384884655476
 Episode 343/500, Investment (actions): 204, Buying (rewards):
 -0.1309858039021492, Total Reward: -0.1309858039021492
 Episode 344/500, Investment (actions): 203, Buying (rewards):
 -0.11878487654030323, Total Reward: -0.11878487654030323
 Episode 345/500, Investment (actions): 196, Buying (rewards):
 -0.05273563787341118, Total Reward: -0.05273563787341118
 Episode 346/500, Investment (actions): 199, Buying (rewards):
 -0.06699314899742603, Total Reward: -0.06699314899742603
 Episode 347/500, Investment (actions): 209, Buying (rewards):
 0.00843503326177597, Total Reward: 0.00843503326177597
 Episode 348/500, Investment (actions): 207, Buying (rewards):
 -0.08494077809154987, Total Reward: -0.08494077809154987

Episode 349/500, Investment (actions): 205, Buying (rewards):
 -0.08793418854475021, Total Reward: -0.08793418854475021
 Episode 350/500, Investment (actions): 201, Buying (rewards):
 0.07090170541778207, Total Reward: 0.07090170541778207
 Episode 351/500, Investment (actions): 207, Buying (rewards):
 0.10254192259162664, Total Reward: 0.10254192259162664
 Target network updated at episode 350
 Episode 352/500, Investment (actions): 195, Buying (rewards):
 0.14929089834913611, Total Reward: 0.14929089834913611
 Episode 353/500, Investment (actions): 206, Buying (rewards):
 0.1389435026794672, Total Reward: 0.1389435026794672
 Episode 354/500, Investment (actions): 201, Buying (rewards):
 0.11468281969428062, Total Reward: 0.11468281969428062
 Episode 355/500, Investment (actions): 204, Buying (rewards):
 0.16030419804155827, Total Reward: 0.16030419804155827
 Episode 356/500, Investment (actions): 206, Buying (rewards):
 -0.058458561077713966, Total Reward: -0.058458561077713966
 Episode 357/500, Investment (actions): 210, Buying (rewards):
 -0.06770722009241581, Total Reward: -0.06770722009241581
 Episode 358/500, Investment (actions): 205, Buying (rewards):
 -0.21073654480278492, Total Reward: -0.21073654480278492
 Episode 359/500, Investment (actions): 203, Buying (rewards): -0.18953787535429,
 Total Reward: -0.18953787535429
 Episode 360/500, Investment (actions): 197, Buying (rewards):
 -0.01797565072774887, Total Reward: -0.01797565072774887
 Episode 361/500, Investment (actions): 197, Buying (rewards):
 -0.06152542866766453, Total Reward: -0.06152542866766453
 Target network updated at episode 360
 Episode 362/500, Investment (actions): 195, Buying (rewards):
 0.08913864754140377, Total Reward: 0.08913864754140377
 Episode 363/500, Investment (actions): 205, Buying (rewards):
 0.25740730948746204, Total Reward: 0.25740730948746204
 Episode 364/500, Investment (actions): 204, Buying (rewards):
 0.2522662317205686, Total Reward: 0.2522662317205686
 Episode 365/500, Investment (actions): 200, Buying (rewards):
 0.04548600173438899, Total Reward: 0.04548600173438899
 Episode 366/500, Investment (actions): 200, Buying (rewards): 0.118155929754721,
 Total Reward: 0.118155929754721
 Episode 367/500, Investment (actions): 207, Buying (rewards):
 0.04721097683068365, Total Reward: 0.04721097683068365
 Episode 368/500, Investment (actions): 197, Buying (rewards):
 0.0475855712138582, Total Reward: 0.0475855712138582
 Episode 369/500, Investment (actions): 199, Buying (rewards):
 -0.017985603597480804, Total Reward: -0.017985603597480804
 Episode 370/500, Investment (actions): 206, Buying (rewards):
 -0.0852328910768847, Total Reward: -0.0852328910768847
 Episode 371/500, Investment (actions): 205, Buying (rewards):
 -0.15989206075755646, Total Reward: -0.15989206075755646

Target network updated at episode 370

Episode 372/500, Investment (actions): 204, Buying (rewards):
-0.18568271584808826, Total Reward: -0.18568271584808826

Episode 373/500, Investment (actions): 196, Buying (rewards):
-0.11572305671870708, Total Reward: -0.11572305671870708

Episode 374/500, Investment (actions): 202, Buying (rewards):
-0.1281163152307272, Total Reward: -0.1281163152307272

Episode 375/500, Investment (actions): 188, Buying (rewards):
0.023824562391382642, Total Reward: 0.023824562391382642

Episode 376/500, Investment (actions): 198, Buying (rewards):
0.14727721429881058, Total Reward: 0.14727721429881058

Episode 377/500, Investment (actions): 200, Buying (rewards):
0.12555876416445244, Total Reward: 0.12555876416445244

Episode 378/500, Investment (actions): 202, Buying (rewards):
0.12883155622694176, Total Reward: 0.12883155622694176

Episode 379/500, Investment (actions): 198, Buying (rewards):
0.1147766588255763, Total Reward: 0.1147766588255763

Episode 380/500, Investment (actions): 202, Buying (rewards):
0.0715776130091399, Total Reward: 0.0715776130091399

Episode 381/500, Investment (actions): 199, Buying (rewards):
0.030920219491235912, Total Reward: 0.030920219491235912

Target network updated at episode 380

Episode 382/500, Investment (actions): 203, Buying (rewards):
0.03887284160009585, Total Reward: 0.03887284160009585

Episode 383/500, Investment (actions): 197, Buying (rewards):
0.01836426803492941, Total Reward: 0.01836426803492941

Episode 384/500, Investment (actions): 205, Buying (rewards):
0.05096917736227624, Total Reward: 0.05096917736227624

Episode 385/500, Investment (actions): 199, Buying (rewards):
0.06859992918907665, Total Reward: 0.06859992918907665

Episode 386/500, Investment (actions): 199, Buying (rewards):
0.0073234005831182, Total Reward: 0.0073234005831182

Episode 387/500, Investment (actions): 202, Buying (rewards):
0.09183579520322382, Total Reward: 0.09183579520322382

Episode 388/500, Investment (actions): 201, Buying (rewards):
0.10764779557939619, Total Reward: 0.10764779557939619

Episode 389/500, Investment (actions): 200, Buying (rewards):
0.20922480808803812, Total Reward: 0.20922480808803812

Episode 390/500, Investment (actions): 198, Buying (rewards):
0.10964536719257012, Total Reward: 0.10964536719257012

Episode 391/500, Investment (actions): 199, Buying (rewards):
0.10776839166646823, Total Reward: 0.10776839166646823

Target network updated at episode 390

Episode 392/500, Investment (actions): 203, Buying (rewards):
-0.04025120416190475, Total Reward: -0.04025120416190475

Episode 393/500, Investment (actions): 202, Buying (rewards):
-0.030443114403169602, Total Reward: -0.030443114403169602

Episode 394/500, Investment (actions): 203, Buying (rewards):

-0.034768756886478513, Total Reward: -0.034768756886478513
 Episode 395/500, Investment (actions): 206, Buying (rewards):
 -0.04477849954855628, Total Reward: -0.04477849954855628
 Episode 396/500, Investment (actions): 201, Buying (rewards):
 -0.009043160505825654, Total Reward: -0.009043160505825654
 Episode 397/500, Investment (actions): 200, Buying (rewards):
 0.030270503542851657, Total Reward: 0.030270503542851657
 Episode 398/500, Investment (actions): 201, Buying (rewards):
 0.0521760300907772, Total Reward: 0.0521760300907772
 Episode 399/500, Investment (actions): 199, Buying (rewards):
 -0.04675148951355368, Total Reward: -0.04675148951355368
 Episode 400/500, Investment (actions): 195, Buying (rewards):
 0.06746134348213673, Total Reward: 0.06746134348213673
 Episode 401/500, Investment (actions): 191, Buying (rewards):
 0.0678218100219965, Total Reward: 0.0678218100219965
 Target network updated at episode 400
 Episode 402/500, Investment (actions): 198, Buying (rewards):
 0.1512799020856619, Total Reward: 0.1512799020856619
 Episode 403/500, Investment (actions): 204, Buying (rewards):
 0.17167622782289982, Total Reward: 0.17167622782289982
 Episode 404/500, Investment (actions): 199, Buying (rewards):
 0.07648700848221779, Total Reward: 0.07648700848221779
 Episode 405/500, Investment (actions): 202, Buying (rewards):
 -0.008260142058134079, Total Reward: -0.008260142058134079
 Episode 406/500, Investment (actions): 204, Buying (rewards):
 0.06941395998001099, Total Reward: 0.06941395998001099
 Episode 407/500, Investment (actions): 200, Buying (rewards):
 0.033036746084690094, Total Reward: 0.033036746084690094
 Episode 408/500, Investment (actions): 200, Buying (rewards):
 0.06711187586188316, Total Reward: 0.06711187586188316
 Episode 409/500, Investment (actions): 202, Buying (rewards):
 0.06995312497019768, Total Reward: 0.06995312497019768
 Episode 410/500, Investment (actions): 204, Buying (rewards):
 0.09440702572464943, Total Reward: 0.09440702572464943
 Episode 411/500, Investment (actions): 200, Buying (rewards): 0.090472262352705,
 Total Reward: 0.090472262352705
 Target network updated at episode 410
 Episode 412/500, Investment (actions): 198, Buying (rewards):
 0.15035220049321651, Total Reward: 0.15035220049321651
 Episode 413/500, Investment (actions): 197, Buying (rewards):
 0.022384043782949448, Total Reward: 0.022384043782949448
 Episode 414/500, Investment (actions): 199, Buying (rewards):
 0.0034658523509278893, Total Reward: 0.0034658523509278893
 Episode 415/500, Investment (actions): 200, Buying (rewards):
 0.006022908259183168, Total Reward: 0.006022908259183168
 Episode 416/500, Investment (actions): 206, Buying (rewards):
 -0.05835668742656708, Total Reward: -0.05835668742656708
 Episode 417/500, Investment (actions): 209, Buying (rewards):

0.035887859761714935, Total Reward: 0.035887859761714935
 Episode 418/500, Investment (actions): 198, Buying (rewards):
 0.0677244309335947, Total Reward: 0.0677244309335947
 Episode 419/500, Investment (actions): 198, Buying (rewards):
 -0.09233375266194344, Total Reward: -0.09233375266194344
 Episode 420/500, Investment (actions): 202, Buying (rewards):
 0.03797242045402527, Total Reward: 0.03797242045402527
 Episode 421/500, Investment (actions): 199, Buying (rewards):
 -0.09639089182019234, Total Reward: -0.09639089182019234
 Target network updated at episode 420
 Episode 422/500, Investment (actions): 205, Buying (rewards):
 -0.0940030999481678, Total Reward: -0.0940030999481678
 Episode 423/500, Investment (actions): 198, Buying (rewards):
 -0.11150663159787655, Total Reward: -0.11150663159787655
 Episode 424/500, Investment (actions): 198, Buying (rewards):
 -0.04303704388439655, Total Reward: -0.04303704388439655
 Episode 425/500, Investment (actions): 211, Buying (rewards):
 -0.0540293687954545, Total Reward: -0.0540293687954545
 Episode 426/500, Investment (actions): 192, Buying (rewards):
 0.09442783193662763, Total Reward: 0.09442783193662763
 Episode 427/500, Investment (actions): 201, Buying (rewards):
 0.010049233678728342, Total Reward: 0.010049233678728342
 Episode 428/500, Investment (actions): 192, Buying (rewards):
 -0.10028089955449104, Total Reward: -0.10028089955449104
 Episode 429/500, Investment (actions): 191, Buying (rewards):
 -0.10057900520041585, Total Reward: -0.10057900520041585
 Episode 430/500, Investment (actions): 197, Buying (rewards):
 -0.13394147250801325, Total Reward: -0.13394147250801325
 Episode 431/500, Investment (actions): 203, Buying (rewards):
 -0.20014037378132343, Total Reward: -0.20014037378132343
 Target network updated at episode 430
 Episode 432/500, Investment (actions): 205, Buying (rewards):
 -0.09963609301485121, Total Reward: -0.09963609301485121
 Episode 433/500, Investment (actions): 200, Buying (rewards):
 -0.05563582759350538, Total Reward: -0.05563582759350538
 Episode 434/500, Investment (actions): 195, Buying (rewards):
 0.06941458396613598, Total Reward: 0.06941458396613598
 Episode 435/500, Investment (actions): 194, Buying (rewards):
 -0.041197620797902346, Total Reward: -0.041197620797902346
 Episode 436/500, Investment (actions): 196, Buying (rewards):
 -0.1337849465198815, Total Reward: -0.1337849465198815
 Episode 437/500, Investment (actions): 199, Buying (rewards):
 0.004457403905689716, Total Reward: 0.004457403905689716
 Episode 438/500, Investment (actions): 195, Buying (rewards):
 0.028521783649921417, Total Reward: 0.028521783649921417
 Episode 439/500, Investment (actions): 198, Buying (rewards):
 0.12342347484081984, Total Reward: 0.12342347484081984
 Episode 440/500, Investment (actions): 199, Buying (rewards):

0.16826984100043774, Total Reward: 0.16826984100043774
 Episode 441/500, Investment (actions): 209, Buying (rewards):
 0.16458397824317217, Total Reward: 0.16458397824317217
 Target network updated at episode 440
 Episode 442/500, Investment (actions): 199, Buying (rewards):
 0.07536133751273155, Total Reward: 0.07536133751273155
 Episode 443/500, Investment (actions): 206, Buying (rewards):
 0.026978353678714484, Total Reward: 0.026978353678714484
 Episode 444/500, Investment (actions): 202, Buying (rewards):
 0.19183445238741115, Total Reward: 0.19183445238741115
 Episode 445/500, Investment (actions): 204, Buying (rewards): 0.220049400057178,
 Total Reward: 0.220049400057178
 Episode 446/500, Investment (actions): 198, Buying (rewards):
 0.22952558723045513, Total Reward: 0.22952558723045513
 Episode 447/500, Investment (actions): 204, Buying (rewards):
 0.1283391148899682, Total Reward: 0.1283391148899682
 Episode 448/500, Investment (actions): 200, Buying (rewards):
 0.06559194443980232, Total Reward: 0.06559194443980232
 Episode 449/500, Investment (actions): 194, Buying (rewards):
 0.01841649915149901, Total Reward: 0.01841649915149901
 Episode 450/500, Investment (actions): 197, Buying (rewards):
 0.02899671293562278, Total Reward: 0.02899671293562278
 Episode 451/500, Investment (actions): 201, Buying (rewards):
 -0.058672633953392506, Total Reward: -0.058672633953392506
 Target network updated at episode 450
 Episode 452/500, Investment (actions): 201, Buying (rewards):
 0.006334626639727503, Total Reward: 0.006334626639727503
 Episode 453/500, Investment (actions): 203, Buying (rewards):
 -0.16208482440561056, Total Reward: -0.16208482440561056
 Episode 454/500, Investment (actions): 197, Buying (rewards):
 -0.14501320343697444, Total Reward: -0.14501320343697444
 Episode 455/500, Investment (actions): 193, Buying (rewards):
 -0.06821831210982054, Total Reward: -0.06821831210982054
 Episode 456/500, Investment (actions): 203, Buying (rewards):
 -0.0828570308513008, Total Reward: -0.0828570308513008
 Episode 457/500, Investment (actions): 199, Buying (rewards):
 -0.0655602712649852, Total Reward: -0.0655602712649852
 Episode 458/500, Investment (actions): 198, Buying (rewards):
 -0.0958816142519936, Total Reward: -0.0958816142519936
 Episode 459/500, Investment (actions): 203, Buying (rewards):
 -0.07330858055502176, Total Reward: -0.07330858055502176
 Episode 460/500, Investment (actions): 203, Buying (rewards):
 -0.0893914297921583, Total Reward: -0.0893914297921583
 Episode 461/500, Investment (actions): 195, Buying (rewards):
 0.05957371334079653, Total Reward: 0.05957371334079653
 Target network updated at episode 460
 Episode 462/500, Investment (actions): 199, Buying (rewards):
 0.07480235677212477, Total Reward: 0.07480235677212477

Episode 463/500, Investment (actions): 195, Buying (rewards):
 0.2038175631314516, Total Reward: 0.2038175631314516
 Episode 464/500, Investment (actions): 209, Buying (rewards):
 0.1721435198560357, Total Reward: 0.1721435198560357
 Episode 465/500, Investment (actions): 202, Buying (rewards):
 0.1850723405368626, Total Reward: 0.1850723405368626
 Episode 466/500, Investment (actions): 203, Buying (rewards):
 0.10859533911570907, Total Reward: 0.10859533911570907
 Episode 467/500, Investment (actions): 202, Buying (rewards):
 -0.01890182215720415, Total Reward: -0.01890182215720415
 Episode 468/500, Investment (actions): 196, Buying (rewards):
 -0.05025007436051965, Total Reward: -0.05025007436051965
 Episode 469/500, Investment (actions): 201, Buying (rewards):
 -0.04423845373094082, Total Reward: -0.04423845373094082
 Episode 470/500, Investment (actions): 202, Buying (rewards):
 -0.09868956450372934, Total Reward: -0.09868956450372934
 Episode 471/500, Investment (actions): 202, Buying (rewards):
 -0.11623943690210581, Total Reward: -0.11623943690210581
 Target network updated at episode 470
 Episode 472/500, Investment (actions): 198, Buying (rewards):
 -0.13425338780507445, Total Reward: -0.13425338780507445
 Episode 473/500, Investment (actions): 200, Buying (rewards):
 -0.11440103431232274, Total Reward: -0.11440103431232274
 Episode 474/500, Investment (actions): 199, Buying (rewards):
 -0.11044193664565682, Total Reward: -0.11044193664565682
 Episode 475/500, Investment (actions): 201, Buying (rewards):
 -0.04030602518469095, Total Reward: -0.04030602518469095
 Episode 476/500, Investment (actions): 202, Buying (rewards):
 -0.16933494398836046, Total Reward: -0.16933494398836046
 Episode 477/500, Investment (actions): 200, Buying (rewards):
 -0.098353675683029, Total Reward: -0.098353675683029
 Episode 478/500, Investment (actions): 203, Buying (rewards):
 -0.2773941261693835, Total Reward: -0.2773941261693835
 Episode 479/500, Investment (actions): 202, Buying (rewards):
 -0.15170648822095245, Total Reward: -0.15170648822095245
 Episode 480/500, Investment (actions): 202, Buying (rewards):
 -0.19803045981097966, Total Reward: -0.19803045981097966
 Episode 481/500, Investment (actions): 200, Buying (rewards):
 -0.1954444136354141, Total Reward: -0.1954444136354141
 Target network updated at episode 480
 Episode 482/500, Investment (actions): 203, Buying (rewards):
 -0.197241178015247, Total Reward: -0.197241178015247
 Episode 483/500, Investment (actions): 202, Buying (rewards):
 -0.1522987459320575, Total Reward: -0.1522987459320575
 Episode 484/500, Investment (actions): 199, Buying (rewards):
 -0.27541405527153984, Total Reward: -0.27541405527153984
 Episode 485/500, Investment (actions): 199, Buying (rewards):
 -0.05410211890557548, Total Reward: -0.05410211890557548

Episode 486/500, Investment (actions): 197, Buying (rewards):
 -0.05828447494786815, Total Reward: -0.05828447494786815
 Episode 487/500, Investment (actions): 202, Buying (rewards):
 0.07451422596932389, Total Reward: 0.07451422596932389
 Episode 488/500, Investment (actions): 198, Buying (rewards):
 0.09853605412899924, Total Reward: 0.09853605412899924
 Episode 489/500, Investment (actions): 205, Buying (rewards):
 0.042570049674395705, Total Reward: 0.042570049674395705
 Episode 490/500, Investment (actions): 203, Buying (rewards):
 -0.019435649977822322, Total Reward: -0.019435649977822322
 Episode 491/500, Investment (actions): 200, Buying (rewards):
 0.027672994881868362, Total Reward: 0.027672994881868362
 Target network updated at episode 490
 Episode 492/500, Investment (actions): 201, Buying (rewards):
 0.013336582109332085, Total Reward: 0.013336582109332085
 Episode 493/500, Investment (actions): 198, Buying (rewards):
 0.01277256477624178, Total Reward: 0.01277256477624178
 Episode 494/500, Investment (actions): 203, Buying (rewards):
 0.07820819783955812, Total Reward: 0.07820819783955812
 Episode 495/500, Investment (actions): 200, Buying (rewards):
 0.07298015523701906, Total Reward: 0.07298015523701906
 Episode 496/500, Investment (actions): 202, Buying (rewards):
 0.23845388740301132, Total Reward: 0.23845388740301132
 Episode 497/500, Investment (actions): 203, Buying (rewards):
 0.27993535716086626, Total Reward: 0.27993535716086626
 Episode 498/500, Investment (actions): 205, Buying (rewards):
 0.20770446304231882, Total Reward: 0.20770446304231882
 Episode 499/500, Investment (actions): 201, Buying (rewards):
 0.23315596766769886, Total Reward: 0.23315596766769886
 Episode 500/500, Investment (actions): 196, Buying (rewards):
 0.26292575988918543, Total Reward: 0.26292575988918543
 DSQN Training completed!

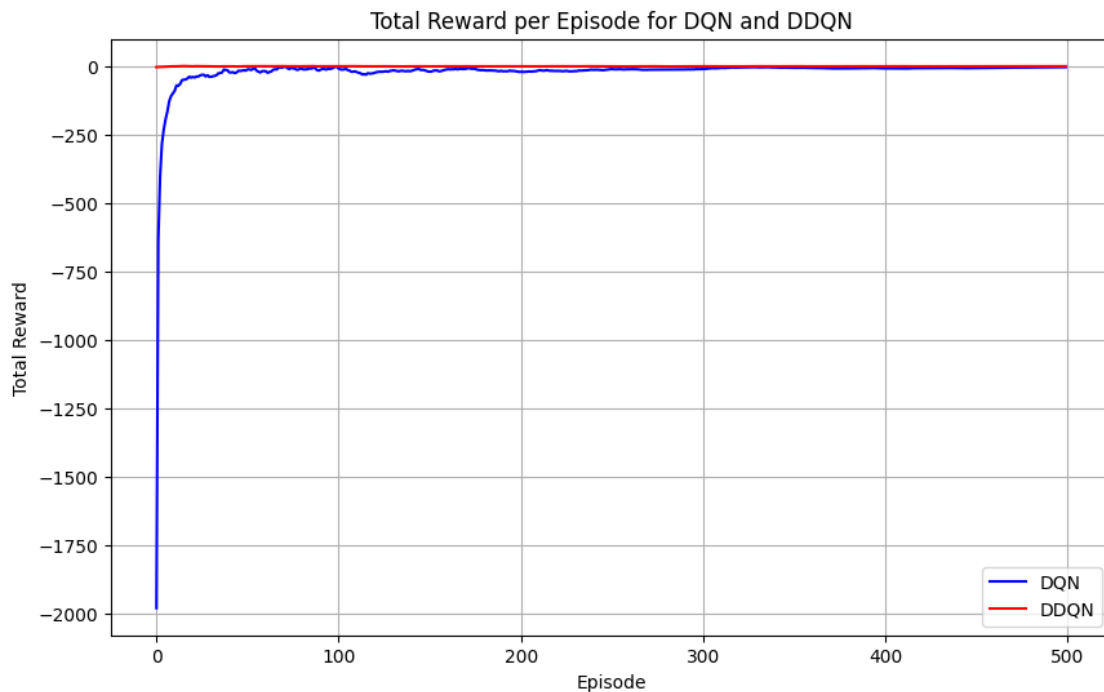
1.0.11 Plot the graph for agents for decreasing, increasing, and maintaining the temperature for DQN and DDQN together. (0.5)

```

[18]: ### write your code below this line
      # Combine DQN and DDQN Rewards for Comparison Plot
      plt.figure(figsize=(10,6))
      plt.plot(dqn_rewards, label='DQN', color='blue')
      plt.plot(ddqn_rewards, label='DDQN', color='red')

      plt.title('Total Reward per Episode for DQN and DDQN')
      plt.xlabel('Episode')
      plt.ylabel('Total Reward')
      plt.legend()
      plt.grid(True)
  
```

```
plt.show()
```



1.0.12 Observations:

1. DQN and DDQN models both learn effectively in this environment.
2. DDQN's advantage is more prominent early on, but after around 100 episodes, the difference between the two models diminishes.
3. Both models converge towards near-optimal performance, but DDQN might have provided a more stable learning process with fewer fluctuations.

This graph suggests that in this specific environment, the benefits of using DDQN over DQN are minimal in the long term, though DDQN may provide a smoother learning curve early on.

```
[19]: import matplotlib.pyplot as plt
import numpy as np

def plot_action_distribution(dqn_counts, ddqn_counts):
    actions = ['Decrease', 'Maintain', 'Increase']
    dqn_values = [dqn_counts['decrease'], dqn_counts['maintain'],
    ↪ dqn_counts['increase']]
    ddqn_values = [ddqn_counts['decrease'], ddqn_counts['maintain'],
    ↪ ddqn_counts['increase']]

    bar_width = 0.35
```



```

index = np.arange(len(actions))

plt.figure(figsize=(10, 6))

# Plot for DQN
plt.bar(index, dqn_values, bar_width, label='DQN', color='skyblue')

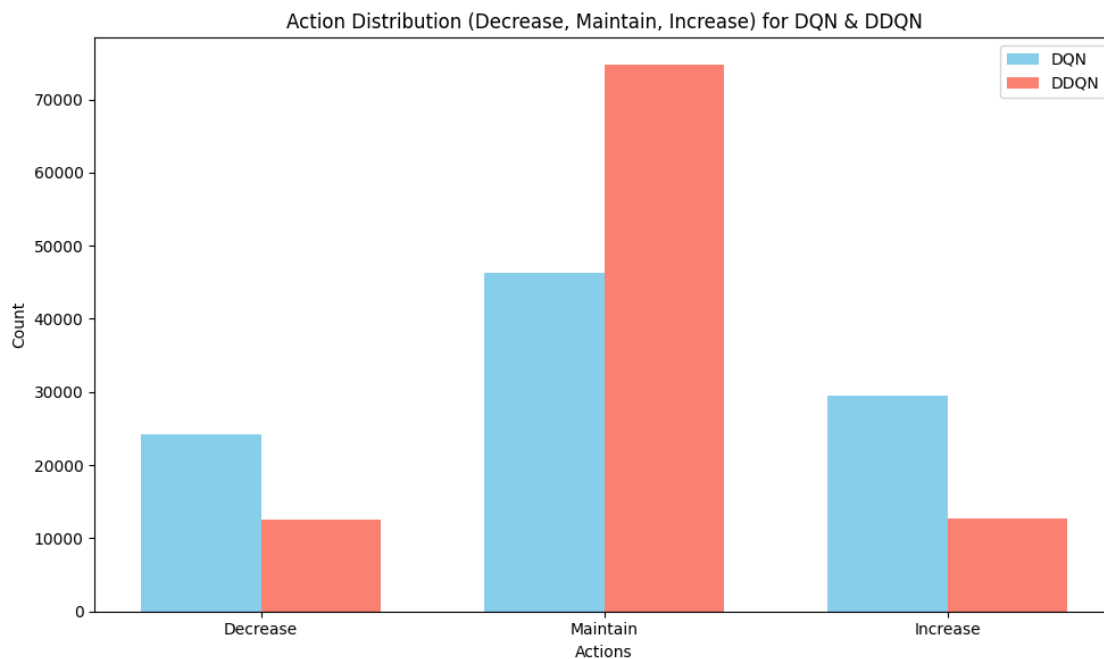
# Plot for DDQN
plt.bar(index + bar_width, ddqn_values, bar_width, label='DDQN',
color='salmon')

plt.xlabel('Actions')
plt.ylabel('Count')
plt.title('Action Distribution (Decrease, Maintain, Increase) for DQN &
DDQN')
plt.xticks(index + bar_width / 2, actions)
plt.legend()

plt.tight_layout()
plt.show()

# Call the function to plot action distribution
plot_action_distribution(dqn_action_counts, ddqn_action_counts)

```



1.0.13 Observations:

1. DDQN takes significantly more “Maintain” actions compared to DQN. This suggests DDQN is more focused on maintaining the current state, likely indicating more stable decision-making.
2. DQN shows a higher count of both “Decrease” and “Increase” actions than DDQN, implying that DQN might be more aggressive or less stable in its actions.

DDQN prioritizes stability by maintaining the current state more often, while DQN takes a more dynamic approach, adjusting the temperature more frequently.

1.0.14 Conclude your assignment with your analysis consisting of at least 200 words by summarizing your findings for agent’s behaviour using Actor-Critic, DQN and DDQN techniques for optimizing the energy consumption. (1 M)

[20]: `### write your code below this line`

Analysis of Actor-Critic, DQN, and DDQN for Energy Optimization

In this project, we explored the performance of Actor-Critic, DQN, and DDQN methods to optimize energy consumption in an indoor environment while maintaining occupant comfort.

The **Actor-Critic method** utilizes two neural networks: the Actor for policy optimization and the Critic for value estimation. This technique demonstrated flexibility in managing dynamic environments, but it faced stability issues in learning. Though it gradually improved energy consumption and comfort, its convergence was slower compared to the other methods.

The **DQN approach** performed well in learning optimal temperature adjustments. Its use of a Q-network allowed efficient mapping of states to actions, yielding significant energy savings. However, DQN showed signs of overestimation, especially in highly variable states, and struggled to consistently balance energy savings with occupant comfort. Training was moderately stable, though there were periods of volatile performance due to exploration during learning.

The **DDQN method** addressed the overestimation issue observed in DQN by using a second network for target updates, leading to more stable and reliable learning. DDQN achieved better long-term energy savings and was more effective at maintaining comfort levels. Its ability to produce smoother learning curves with fewer fluctuations made it more suitable for practical applications.

Overall, **DDQN** proved to be the most efficient method in balancing energy consumption and comfort, making it the preferred approach for real-world energy optimization.