import random

# Kazanç matrisi

payoff\_matrix = [[0, -1, 1], [1, 0, -1], [-1, 1, 0]]

# Q-learning algoritması

Q = [[random.random() \* 0.1 for \_ in range(3)] for \_ in range(3)]

learning\_rate = 0.1 # Öğrenme hızı

discount\_factor = 0.9 # İndirim faktörü

num\_episodes = 1000 # Toplam bölüm sayısı

# Q-learning algoritmasının uygulanması

for episode in range(num\_episodes):

state = random.randint(0, 2) # Rastgele bir başlangıç durumu seç

while True:

action = max(range(3), key=lambda x: Q[state][x]) # En yüksek Q değerine sahip aksiyonu seç

next\_state = random.choice([0, 1, 2]) # Rastgele bir sonraki durum seç

reward = payoff\_matrix[state][action] # Ödül hesapla

Q[state][action] += learning\_rate \* (reward + discount\_factor \*

max(Q[next\_state]) - Q[state][action]) # Q değerini güncelle

state = next\_state

if state == 0: # Oyunu tamamladığımızda döngüyü sonlandır

break

# Eğitim sonucunda elde edilen Q değerleri

print("Q Değerleri:")

for row in Q:

print(row)