

## **Kelompok 11**

### **Nama Anggota :**

- Rizki Nofrian Wahyudi (Atlas)
- Caesar Calendo Sumarga (Visioner)
- Fadhila Tsani Noor Aisyah (Persevere)
- Steven Dharmawan (Better)

### **Soal :**

1. Cari satu kasus di environment OpenAI yang bisa diselesaikan dengan Monte Carlo, pelajari programnya.
2. Jelaskan yang anda pahami tentang mekanisme kerja dari program tersebut!

Link dalam materi: <http://rl-lab.com/gridworld-mc>

### **Jawab :**

1. Kasus di environment OpenAI yang bisa diselesaikan dengan Monte Carlo, yaitu kasus penyelesaian OpenAI Blackjack Gym Environment menggunakan very-visit MC learning method dengan constant alpha + epsilon greedy strategy method.
2. Mekanisme kerja dari program :

Permainan blackjack merupakan salah satu kasus yang mengimplementasikan reinforcement learning. Reinforcement learning pada permainan black jack dapat menggunakan metode monte carlo. Environment yang digunakan pada program ini merupakan environment dari openAI yaitu "Blackjack-v0". Berikut merupakan penjelasan dari environment "Blackjack-v0":

#### **Menang:**

- Jumlah kartu milik player mendekati angka 21 dan melebihi jumlah kartu milik dealer tanpa melebihi angka 21.

#### **Kalah:**

- Jumlah kartu milik dealer lebih mendekati angka 21 dibandingkan player.
- Jumlah kartu milik player melebihi angka 21.

#### **Peraturan:**

- Kartu jack, queen, dan king memiliki nilai 10.
- Ace bernilai 1 atau 11 menyesuaikan kondisi mana yang lebih baik.
- Player dan dealer dibagikan dua kartu.
- Dealer hanya menunjukkan satu kartu.
- Usable Ace → Jika player mempunyai ace yang dapat dihitung sebagai 11 tanpa melebihi 21.

#### **Action:**

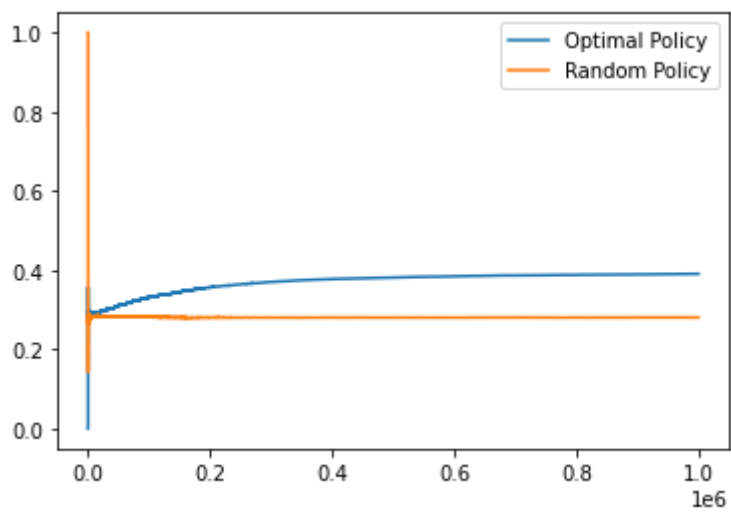
- Hit → Menambah kartu.
- Stick → Tetap dengan kartu saat ini.

Dalam program tersebut, yang dapat kita pahami mengenai langkah-langkah mekanisme kerja dari program tersebut dalam pembuatan program blackjack tersebut, yaitu :

- Mengimport beberapa library yang dibutuhkan
- Mendefinisikan environment + menyiapkan hyperparameters nya
- Menyiapkan arbitrary Q\_function untuk semua states dan actions
- Menghasilkan n-episode
- Mendefinisikan epsilon + melurulkannya dengan konstanta untuk setiap episode
- Mengekstraksi states, actions, rewards dari episode
- Mengulang langkah waktu di setiap episode
- Memperbarui q\_values menggunakan aturan pembaruan MC
- Mengembalikan the policy dan Q function
- Memvisualisasikan hasil

Kesimpulan dari program untuk penyelesaian Open AI blackjack gym environment menggunakan Monte carlo, untuk winrate dari program tersebut ada di sekitar 42% dengan jumlah episode 1000000 dan ketika mendapatkan kartu as sistem lebih banyak melakukan hit dari pada stick begitu pula sebaliknya.

Plot winrate :



Plot policy :

