3. Buat tabel perbandingan dari metoda Dynamic Programming, Monte Carlo, SARSA, Q-Learning, dan Deep Q Learning.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **NO** | **Dynamic Programming** | **Monte Carlo** | **SARSA,** | **Q-Learning,** | **Deep Q Learning** |
| **1.** | metode pemecahan masalah dengan menguraikan solusi menjadi sekumpulan langkah (*step*) atau tahapan (*stage*) sedemikian sehingga solusi dari persoalan dapat dipandang dari serangkaian keputusan yang saling berkaitan. | percobaan pada unsur peluang (atau bersifat probabilistik) dengan menggunakan pengambilan sampel secara acak. | Setiap step terdapat beberapa action yang akan dilakukan, selanjutnya agent akan mengalami pergeseran setelah melakukan action dan mendapatkan reward sesuai dengan action yang dilakukan. Reward tersebut membantu untuk memutuskan action selanjutnya. | pencarian jalur | Deep Q  -  Learning  merupakan salah satu  basis algoritme RL dengan paradigma  model  free  . |
| **2.** | Terdapat sejumlah berhingga pilihan yang mungkin. | Dalam metode monte carlo faktor random dimasukkan kedalam model dengan melibatkan satu atau lebih variable random |  | Q-learning adalah pengembangan dari TemporalDifference yang juga biasa dikenal off-policy TD control. | Deep pada  quadcopter  membutuhkan karakteristik simulator yang  hampir mendekati dunia nyata, memiliki  kemudahan perangkat lunak dan integrasi  perangkat keras. |
| **3** | Solusi pada setiap tahap dibangun dari hasil solusi tahap sebelumnya. |  |  | Semakin kecil nilai γ, maka agen akan semakin mementingkan reward dekat, bukan reward di masa depan. | Deep Q  -  Learning dapat secara efisien mengolah  informasi citra  depth image  yang digunakan  untuk menghindari rintangan. |
| **4.** | Menggunakan persyaratan optimasi dan kendala untuk membatasi sejumlah pilihan yang harus dipertimbangkan pada suatu tahap. |  |  |  | Deep Learning  pada Q  -  Learning bertugas untuk menggantikan  posisi Q  -  Table |
| **5.** | algoritma berbasis *deterministic* | algoritma yang berbasis pendekatan probabilistic |  | Proses pembelajaran Qlearning diawali dengan menginisialisasi nilai action-value function Q(S,A) dan proses perulangan pemilihan aksi A serta nilai action-value function diperbaharui sampai kondisi pembelajaran yang digunakan terpenuhi. | Deep  dapat dipercepat dengan mengubah  hyperparameter  seperti  learning rate  (laju  pembelajaran) atau  batch size  (seberapa banyak  data training dalam satu kali  epoch  ). |