

ABSTRAK

Reinforcement Learning merupakan pendekatan pembelajaran mesin yang digunakan untuk melatih agent melalui interaksi dengan lingkungan, di mana salah satu algoritma yang umum digunakan adalah Q-Learning. Namun, algoritma ini memiliki keterbatasan dalam hal stabilitas pembelajaran dan fluktuasi reward selama proses training. Penelitian ini bertujuan untuk mengoptimalkan algoritma Q-Learning serta menganalisis peningkatan performa agent berdasarkan parameter evaluasi berupa pertumbuhan reward dan stabilitas pembelajaran. Evaluasi dilakukan dengan membandingkan Q-Learning biasa dan Q-Learning yang telah dioptimalkan pada lingkungan permainan ViZDoom dengan skenario Deathmatch dan Defend the Center, menggunakan jumlah episode pelatihan sebanyak 1.000 dan 10.000 episode. Hasil penelitian menunjukkan bahwa optimisasi algoritma Q-Learning mampu meningkatkan pertumbuhan reward dan reward tertinggi secara signifikan pada seluruh skenario pengujian, terutama pada jumlah episode pelatihan yang lebih besar, namun peningkatan tersebut diikuti oleh nilai standar deviasi reward yang lebih tinggi, yang menunjukkan bahwa stabilitas pembelajaran belum sepenuhnya optimal.

Kata kunci: Q-Learning, Double Q-Learning, Reinforcement Learning, ViZDoom, First-Person Shooter.

ABSTRACT

Reinforcement Learning is a machine learning approach used to train agents through interaction with the environment. One commonly used algorithm is Q-Learning. However, this algorithm has limitations in terms of learning stability and reward fluctuations during the training process. This study aims to optimize the Q-Learning algorithm and analyze agent performance improvements based on evaluation parameters such as reward growth and learning stability. The evaluation was conducted by comparing standard Q-Learning and optimized Q-Learning in the ViZDoom game environment with Deathmatch and Defend the Center scenarios, using 1,000 and 10,000 training episodes, respectively. The results show that the Q-Learning algorithm optimization significantly increased reward growth and the highest reward across all test scenarios, especially with a larger number of training episodes. However, this increase was accompanied by a higher reward standard deviation, indicating that learning stability was not fully optimized.

Keywords: Q-Learning, Double Q-Learning, Reinforcement Learning, ViZDoom, First-Person Shooter.