

# The Effect of Discounting Actor-loss in Actor-Critic Algorithm

Jordi Yaputra<sup>1</sup>, Suyanto Suyanto<sup>2</sup>

<sup>1,2</sup> Informatics Faculty, Telkom University, Bandung

[jordiyaputra@student.telkomuniversity.ac.id](mailto:jordiyaputra@student.telkomuniversity.ac.id), [suyanto@telkomuniversity.ac.id](mailto:suyanto@telkomuniversity.ac.id)

---

## Abstract

We analyze and present an experimental approach to see the effect of limiting the Temporal Difference (TD) error in estimating actor-loss on an actor-critic-based agent. The limitation is done by reducing the loss value of an actor to the factor of an epsilon  $\epsilon$  constant. In this experiment, we chose four epsilon values, i.e., 0.01, 0.1, 0.5, and 1.0, where 1.0 means no discount at all. In the experiment, we spawn four agents to solve a trivial task for humans in a custom lightweight Windows Operating System (OS)-like simulation. Each agent receives inputs of the simulation's screen image and controls the cursor inside the simulation to reach for any rendered red circles. After 50 episodes, 50,000 steps in total, each agent achieved about the same success rate with slight differences. The agent given an epsilon value of 0.01 achieved the highest success rate, higher than one without discount learning (epsilon=1.0), although not much.

**Index Terms**—Reinforcement Learning, Actor-Critic, Temporal Difference Learning, Convolutional Neural Network, Artificial Intelligence