**TD3**

**Initial:**

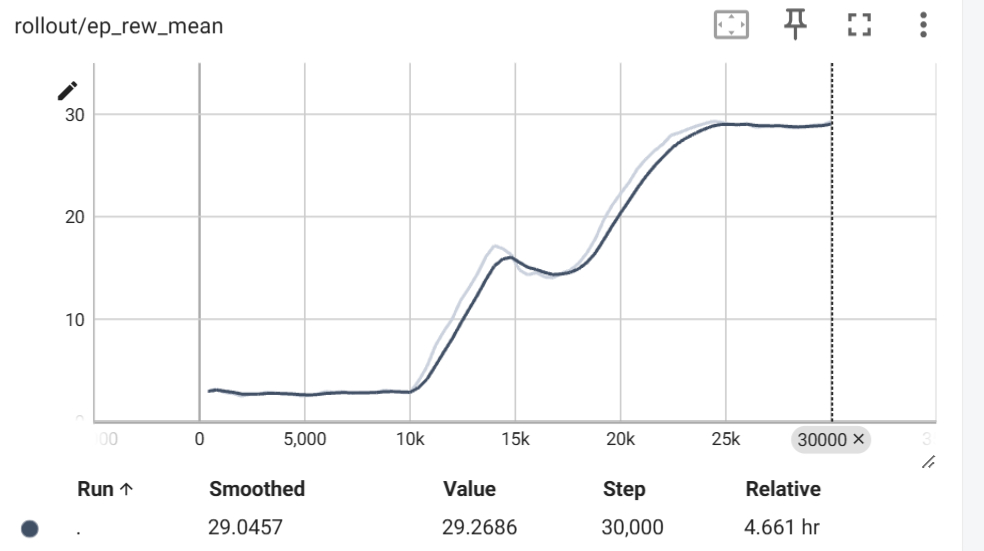
**A graph with numbers and lines

Description automatically generated**

During the initial training phase with **25,000 timesteps**, the model exhibited promising performance. The average reward, as visualized in TensorBoard, showed a steady upward trend. This indicated the potential for further improvement with additional training.

**Action taken:**

To explore whether additional training would enhance performance, I increased the total training timesteps from **25,000** to **35,000**. The decision was based on the observation that the model was still learning and had not yet reached saturation at the initial timestep threshold.



**Results**

After extending the training to **35,000 timesteps**, the following changes were observed:

* The average reward per episode increased noticeably compared to the model trained with 25,000 timesteps.
* The reward trend in TensorBoard confirmed that the model continued to improve with additional training steps, further optimizing its policy for the highway-v0 environment.