

How to evaluate your model and deploy it to AWS DeepRacer for a real-world autonomous experience

In this guide we will look at methods available in AWS DeepRacer to determine how well your model is performing. We will also show you how you can deploy your trained model to the AWS DeepRacer for a real-world racing experience.

The main ways to evaluate model performance are:

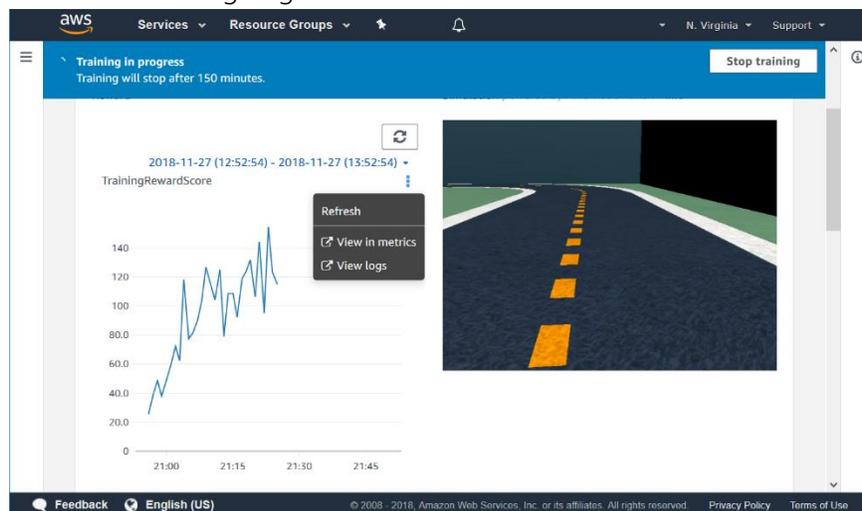
- During training
- Running an evaluation on the model post training
- Deploying the model to the AWS DeepRacer and observing model behavior in the real-world

Assessing model performance during training

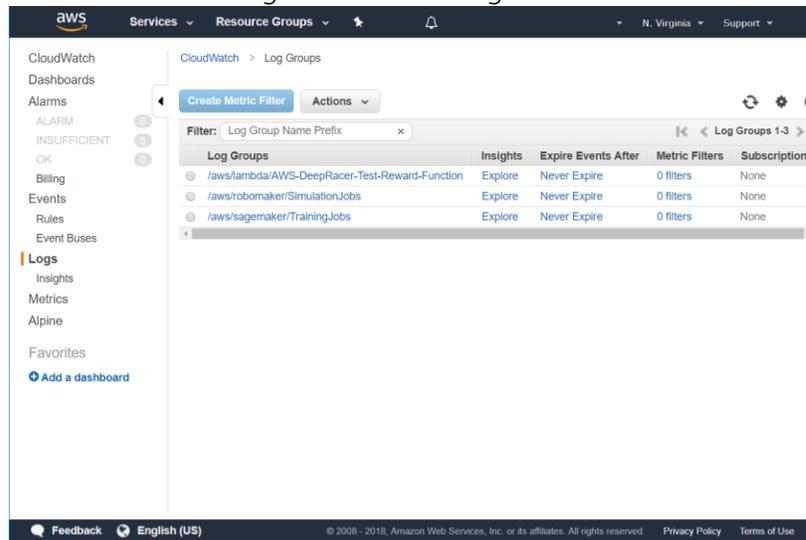
There are a number of ways to assess how well your model is training during model training.

1. You can keep track of training progress by observing the driving behavior of the car in the simulator and by looking at the TrainingRewardScore graph. Initially your car will drive off the track a lot, from pretty much any state, but over time it will stay on track for longer periods as the model trains. Once your model can complete one to two laps without fail, you should consider stopping the training. Training further may result in you overfitting to the simulator, which may result in a poor real-world experience. The TrainingRewardScore graph shows the cumulative reward received per training episode. You can expect this graph to be volatile initially, but over time the graph should trend upwards and to the right and as your model starts converging the average may flatten out. It is always good to try and visually inspect the driving behavior, either during training or using the AWS DeepRacer.
2. You can also inspect the Amazon SageMaker and AWS RoboMaker logs saved in Amazon CloudWatch. To access the logs you can select the 3 dots in the top right of your TrainingRewardScore graph, and select view logs or view in metrics. View logs will show you logs of the output from the Amazon SageMaker and AWS RoboMaker containers. They will contain more detailed information about your training job.

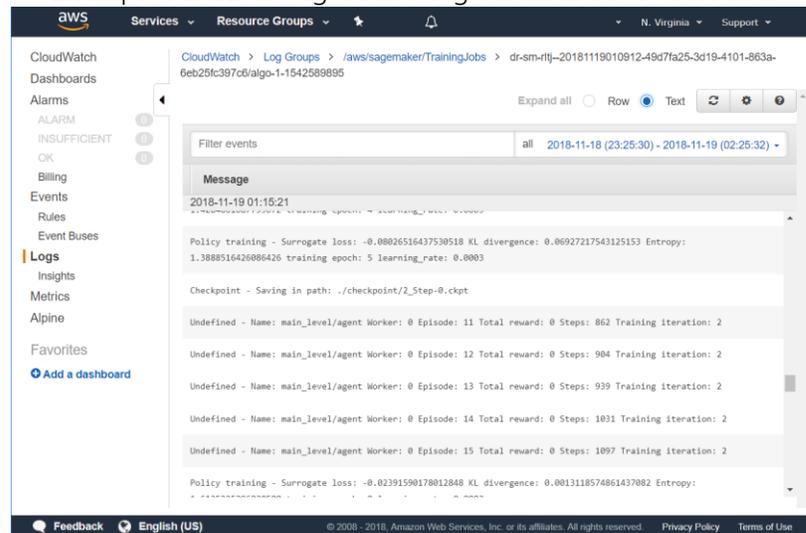
Accessing Logs and Metrics in Amazon CloudWatch



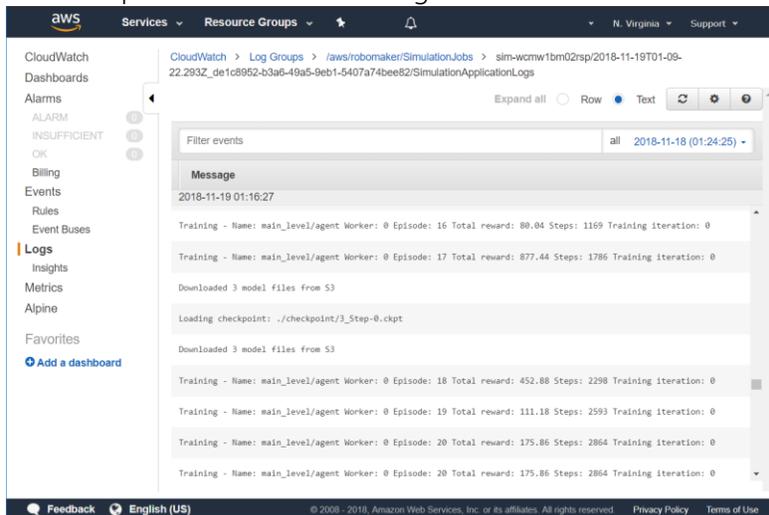
Amazon CloudWatch Logs for Amazon SageMaker and AWS RoboMaker



Example Amazon SageMaker log in Amazon CloudWatch



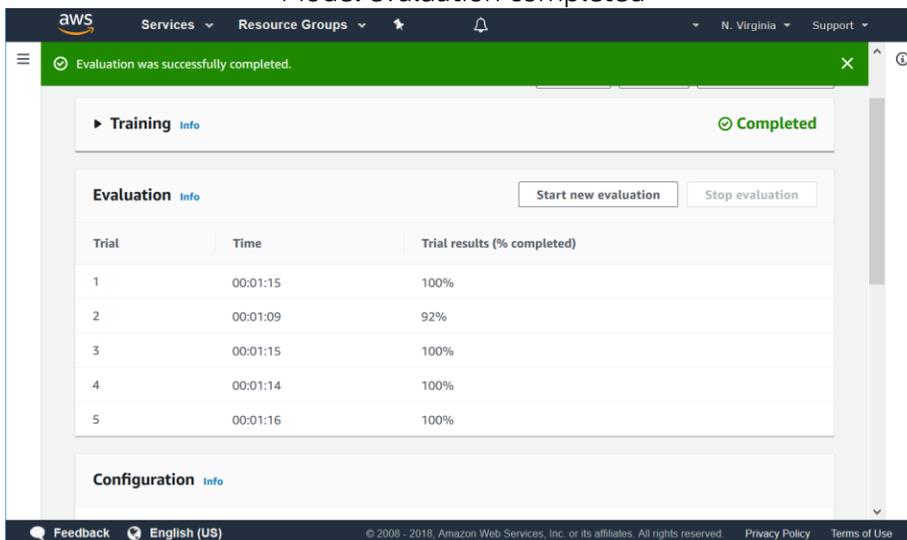
Example AWS RoboMaker log in Amazon CloudWatch



Running model evaluation on trained models

After your model has successfully trained you can evaluate the performance of your model in the simulator. To do so select your model from the list shown in Reinforcement learning. If your model has finished training you should be able to choose **Start Evaluation**. You can now select the track on which you want to evaluate your model and how many laps you want to complete. Once evaluation has been completed you will see the lap time and the percentage of the track completed for each lap. If your model struggles to complete the lap, consider training it some more or changing your reward function.

Model evaluation completed



Deploying the model to the AWS DeepRacer

A more exhilarating way to test your model is by deploying it to your AWS DeepRacer. Please ensure you have completed the “Getting Started” guide that can be found at <https://aws.amazon.com/deepracer/getting-started> to set up your AWS DeepRacer. The guide also provides the steps needed to deploy a trained model to your AWS DeepRacer. At this point please create your racetrack, following the guidelines provided in the Developer Documentation <https://docs.aws.amazon.com/deepracer/latest/developerguide/>.

You can now log into your AWS DeepRacer GUI using the steps outlined in the “Getting Started” guide, and use your model to enable autonomous mode with the vehicle on your track. It is a good idea to start the throttle at 35% and slowly increase it to prevent your car from suddenly taking off and crashing into the barrier you created around your track. As your car battery runs out, the amount of throttle you have to provide may need to increase. A minimum amount of throttle is needed to overcome the initial resistance in the car. Once you have found the sweet spot enjoy experience and proceed to inspect the driving behavior. At this point it is important to call out that there is potentially a substantial difference between the simulator and your real-world race track. We will release new tutorials in future on methods that can be used to overcome this simulation-to-real domain transfer gap.

AWS DeepRacer GUI

