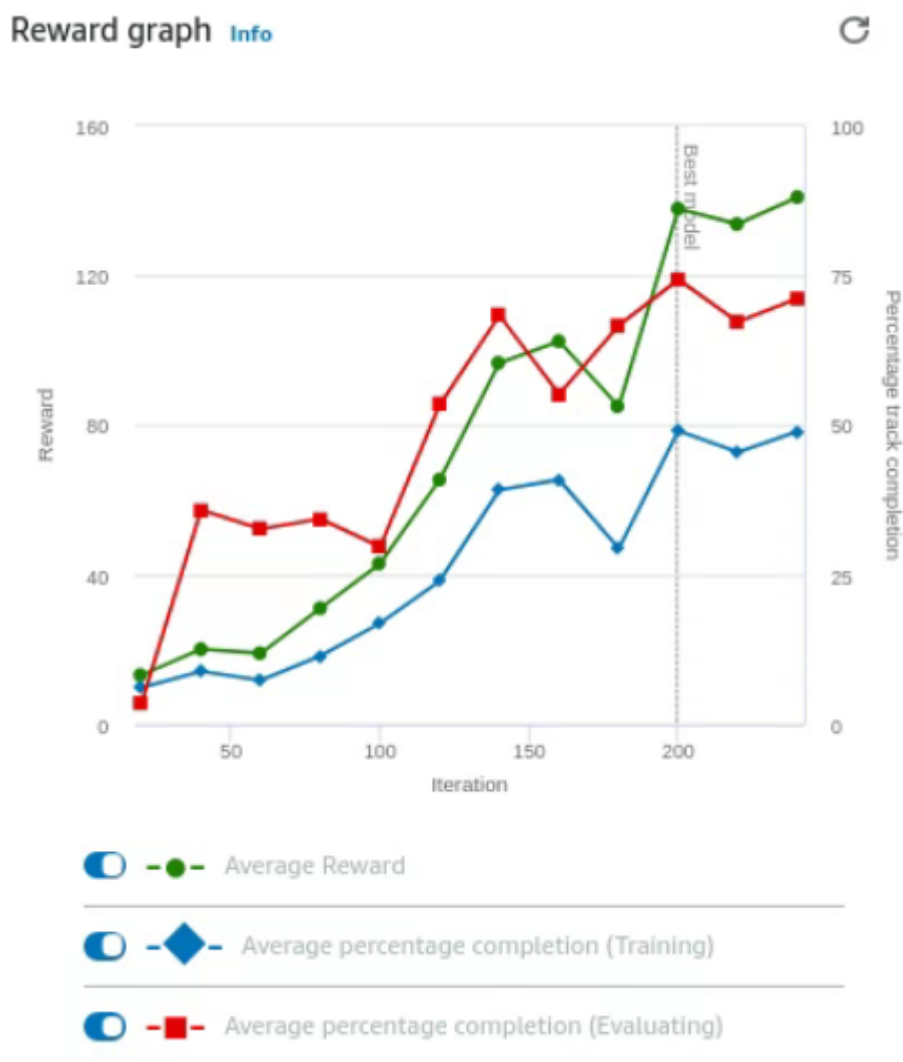


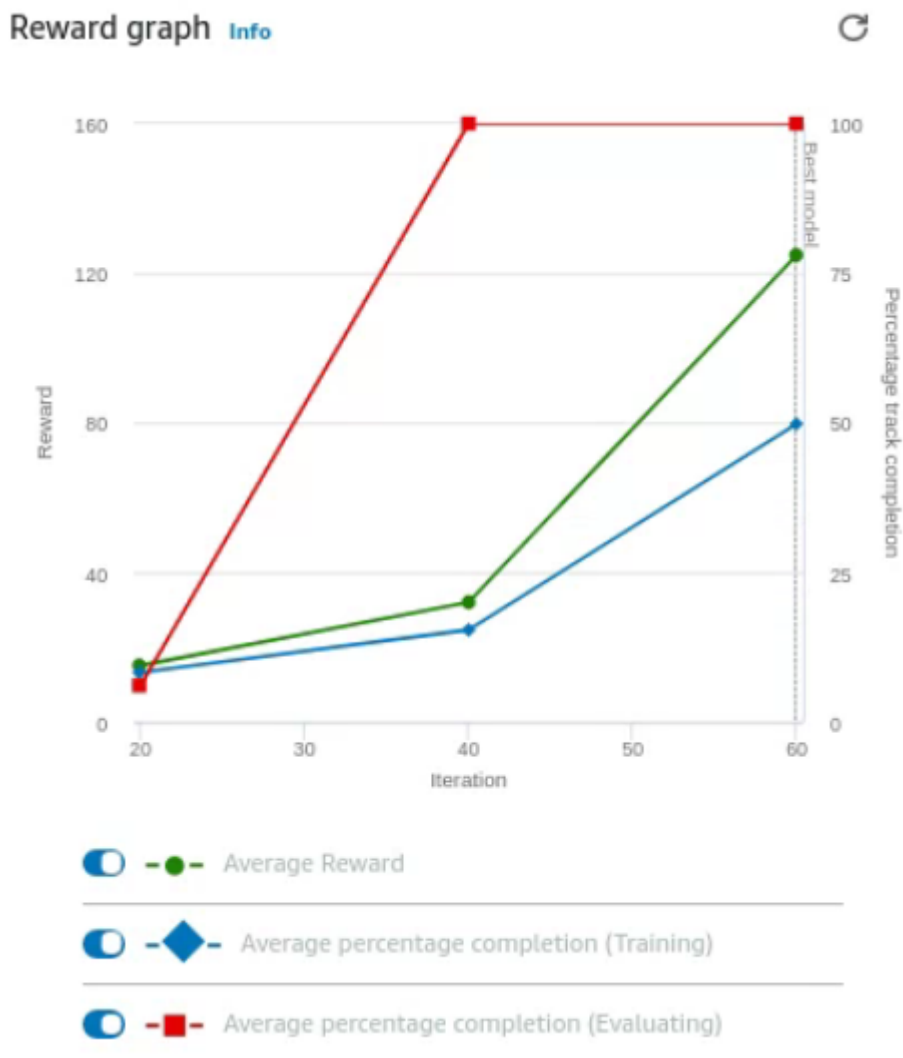
Reward Function 2 vs Reward Function3

이번 page에서는 그래프를 보고 어떤 차이가 있는 지 살펴보려 한다. 이를 통해서 우리가 어떤 방향으로 개발을 진행해야 할지 알 수 있을 것이라 생각한다. 해당 모델들은 똑같은 Track에서 똑같은 1시간이라는 훈련시간을 가지고 진행하였다. 코드는 모두 Reward Function에 기록해놓았다.

Reward Function2의 그래프



Reward Function3의 그래프



두 그래프를 보게 되면 눈에 띄게 Reward Function3가 훨씬 좋은 결과를 보인 것을 알 수 있다. Reward Function3의 경우에는 훨씬 더 적은 iteration으로 track completion을 100%를 달성했을 뿐 아니라, 모델의 학습도 선형적으로 증가하는 모습을 보여주고 있다. 그리고 한번 트랙 완주를 한후, 뚜렷하게 지속적으로 track-off가 없는 주행을 하고 있음을 확인할 수 있다. 그렇다면 무조건 Reward Function3가 더 좋은 성적을 냈을 까? 그건 아니다.

Reward Function2의 기록

Evaluation results

Trial	Time	Trial results (% track completed)	Status
1	00:00:27.879	100%	Lap complete
2	00:00:29.974	100%	Lap complete
3	00:00:30.360	100%	Lap complete

Reward Function3의 기록

Evaluation results

Trial	Time	Trial results (% track completed)	Status
1	00:00:31.483	100%	Lap complete
2	00:00:31.731	100%	Lap complete
3	00:00:30.963	100%	Lap complete

기록에서 알 수 있듯이, 오히려 최고 기록은 Reward Function2의 값이 더 높음을 알 수 있다. 약 3초의 차이가 나고 있다. 이를 통해 우리는 더 짧은 학습이 더 좋은 결과를 내는 것을 아니라는 것을 알아야하며, 그렇다고 학습시간을 너무 늘이는 것 또한 좋지 않다는 것을 알아야한다.(우선, 학습을 오래할 수록 금액이 부과될 것이다.)