

Introduction

- **Machine learning** is the training of a model on data sets before it is used on a test data set to generate a predicted result
- **Reinforcement learning** uses a reward function to encourage certain behaviors
 - If a model adheres to the coded guidelines, then new models can be built off of it to explore other behaviors (Qiang & Zhongli, 2011, p. 1143)
 - Pre-training a model can decrease training time by reducing the initial learning curve (Kim, Cha, Ryu, & Jo, 2019, p. 2)
 - Using past data and trends is practical for the processing of large amounts of signals (Moon, Cheong, Yeom, & Woo, 2019, p. 345)

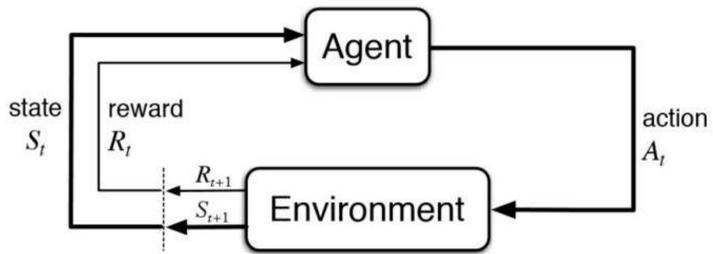


Figure 1. The stages of reinforcement learning
Image Credit: <https://www.kdnuggets.com/2018/03/5-things-reinforcement-learning.html>

- **Amazon Web Services' DeepRacer**
 - Virtual car undergoes training for user-inputted time and then evaluation

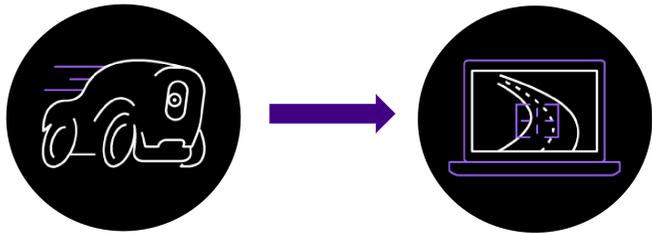


Figure 2. Deep Racer car
Image Credit: <https://aws.amazon.com/deepracer/>
Figure 3. Track visualization
Image Credit: <https://aws.amazon.com/deepracer/>

Research Objectives

- Determining an optimal combination of parameters and their respective rewards in the model
- Building off of previous models to determine how past iterations can influence future performance
- Observe applicability of model on track segments

Methods

Strategies

1. **Track curvature**
 - Relative Distance
$$d = \frac{|w_1 w_3|}{|w_1 w_2| + |w_2 w_3|}$$
 - Bezier Curves
$$K = \frac{|x' y'' - x'' y'|}{((x')^2 + (y')^2)^{3/2}}$$
 - Vector Calculations
$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$
2. **Progress per step**
3. **Speed**

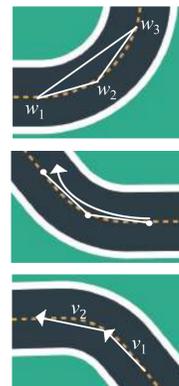


Figure 4. Curvature representations

Development Cycle

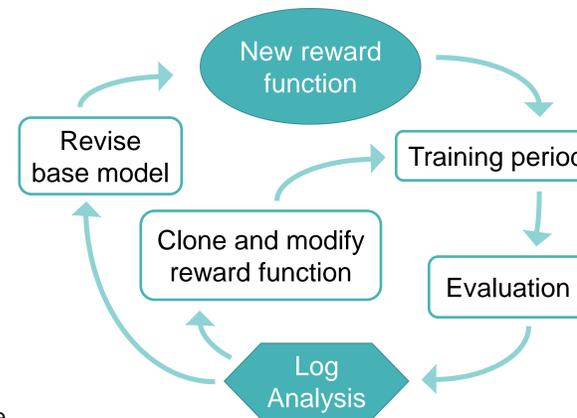


Figure 5. Model development cycle

Results

Figure 6. Distribution of times for evaluation laps in various model series

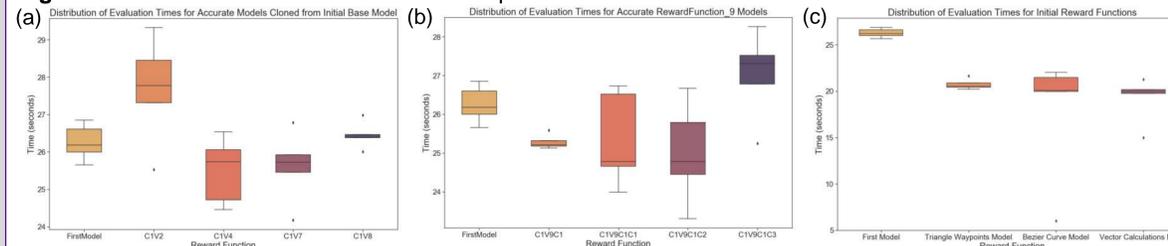


Figure 7. Track paths taken by virtual car during training period for various initial reward functions

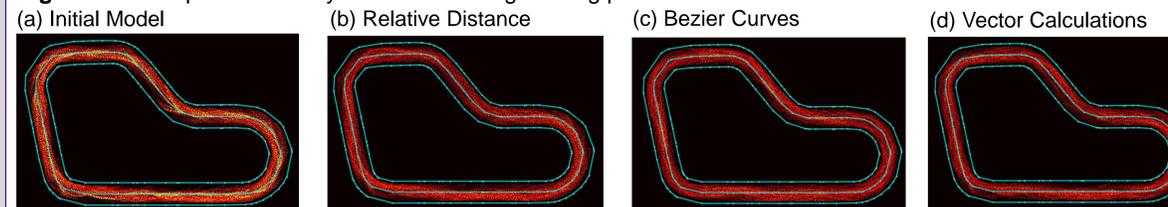


Figure 8. Representation of path taken for an evaluation lap for various initial reward functions

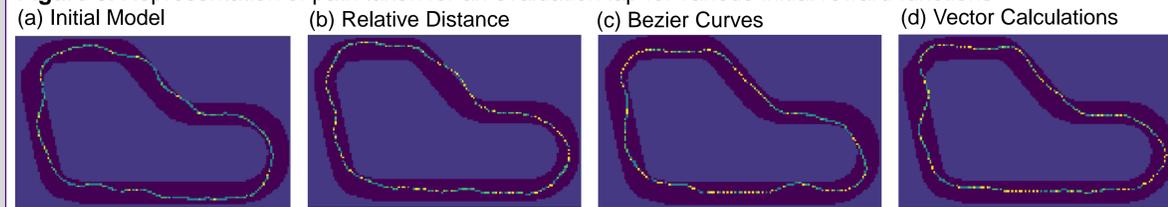
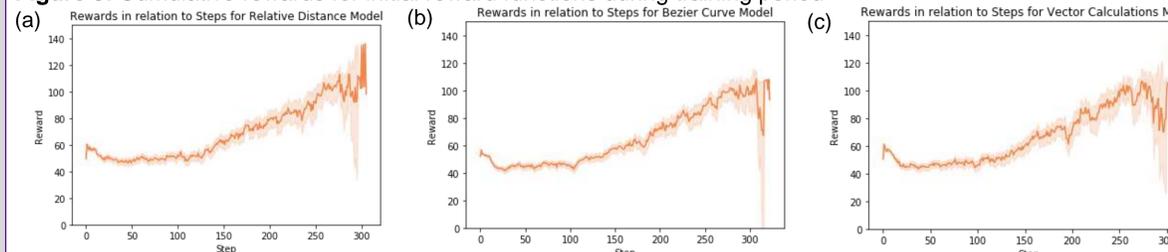


Figure 9. Cumulative rewards for initial reward functions during training period



Conclusion

- **Vector calculations** for curvature yielded the fastest and most accurate performance
- Rewarding progress relative to steps was effective
- Models developed as **newly initialized reward functions** performed better than corresponding cloned models
- Successive iterations of reward functions did not always produce improved models

Future Directions

- Build off of successful models to optimize car path and speed in both virtual and physical Deep Racer
- Explore the use of iterative learning models in navigation based solely on sensor input
- Apply reinforcement learning reward functions to situations where certain behaviors are favored



Figure 10. Deep Racer on a physical track



Figure 11. Interior components of the car



DeepRacer GitHub repository

Acknowledgements

Thank you to my mentor Dr. Denise Szecsei, the Belin-Blank Honors Center, and SSTP for this opportunity. An additional thanks to the University of Iowa's Computer Science Department, the Information Technology Services, and the AWS DeepRacer team for their assistance throughout this project.

References

- Kim, J., Cha, S., Ryu, M. & Jo, M. (2019). Pre-training framework for improving learning speed of reinforcement learning based autonomous vehicles. *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, Auckland, New Zealand, 1-2. doi: 10.23919/ELINFOCOM.2019.8706441.
- Moon, J., Cheong, M., Yeom, I. & Woo, H. (2019). Deep reinforcement learning based sensor data management for vehicles. *2019 International Conference on Information Networking (ICOIN)*, Kuala Lumpur, Malaysia, 345-349. doi: 10.1109/ICOIN.2019.8718108.
- Qiang, W. & Zhongli, Z. (2011). Reinforcement learning model, algorithms and its application. *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, Jilin, China, 1143-1146. doi: 10.1109/MEC.2011.6025669.