Aviral Kumar

Title: The "Deep" in Deep RL: Capacity Loss, and How to Mitigate it in Value-Based RL

Abstract: Despite impressive results from combining ideas from reinforcement learning with deep neural networks, deep reinforcement learning methods appear harder to use in practice onto new problems. One case where this hardness is very apparent is when attempting to scale up deep reinforcement learning methods to use larger neural network models. In particular, while it is pretty straightforward to scale up model capacity in supervised learning, but even instantiations of RL that attempt to learn from static data alone (i.e., offline RL) are not quite easy to utilize larger models. In this talk, I will present some of our work in understanding this challenge by investigating representations learned by the neural networks trained via value-based RL. Specifically, we find that training big models using value-based RL leads to a form of a "capacity loss", where the representations learned by the model gradually lose their ability of distinguishing different inputs as more training is performed. This implicitly reduces the capability of the value network in representing complex value functions, resulting in poor performance over training and a variety of instabilities. We will formalize this empirical observation using the theory of implicit regularization from supervised learning and then show how despite the supposedly favorable benefits of implicit regularization with overparameterized in supervised learning, the very same phenomenon can explain this issue of capacity loss in deep RL. Finally, using these insights, we will build a practical method to mitigate capacity loss and show how this allows us to effectively scale deep RL to use large neural networks, especially when training on diverse data.