# Notes on offline reinforcement learning

Compiled by D. Gueorguiev 1/14/24

## Introductory Notes and Problem Statement

The Reinforcement learning algorithms provide a fundamentally *online* learning paradigm which is one of the biggest obstacles to their adoption in various use cases. The process of reinforcement learning involves iteratively collecting experience by interacting with the environment, typically with the latest learned policy, and the using that experience to improve the policy. In many settings, the online interaction is not really useful or even applicable, because data collection is difficult to guarantee in reliable manner. Furthermore, even in domains where online interaction is feasible, we might still prefer to utilize previously collected data – for example if the domain is with large number of dimensions and complex search space so that effective generalization would require large datasets.

## References

[Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems, Sergey Levine, Aviral Kumar, George Tucker, Justin Fu, 2020](https://github.com/dimitarpg13/self_supervised_learning/blob/main/literature/OfflineReinforcementLearningTutorialReviewLevine2020.pdf)