Paper summary: Loss is its own reward: self-supervision for reinforcement learning

May 30, 2022

1 Idea in few sentances

2 Explanation of the central concept

3 Methodology

4 Initial rambly notes

4.1 Abstract

Using only expected return is weak, can be remedied by auxiliary loss.

4.2 Introduction

If you train an agent and save just the feature extracting portion of the network and use that to retrain the agent, the agent learns faster. Thus having feature extraction is better than not having it. Also, while reward may be sparse, self-supervised auxiliary losses aren't. You can go for discriminative (self-supervised) or generative (unsupervised) learning objectives.

4.3 Method

4.4 Other stuff