

# 基于k最近邻训练的深度Q学习

2018年3月13日

## 1 问题描述

利用深度神经网络解决增强学习问题时，往往存在着这些问题：（1）模型需要通过激励信号学习，估计agent在一段时间内的总激励，而通常激励信号具有稀疏、有噪声和延迟的特点，即采取动作和产生激励信号之间可能间隔多步，与典型的监督学习不同；（2）深度神经网络的预设条件是样本数据相互独立，增强学习问题中的样本数据不独立，特别是相邻序列产生的状态之间高度相关；（3）增强学习中，样本数据的分布随着采取动作的不同而发生变化，深度神经网络则假设数据样本的分布不变。

## 2 算法

## 3 实验结果