Algorithm 1 Least Squares Temporal Differences (LSTD) Deep Q-Learning with Nonlinear Feature Extraction

---

1: Initialize Q-network weights w
2: Initialize feature extraction network weights θ
3: Initialize replay buffer D
4: **for** $episode = 1$ to N **do**
5:      Initialize state S ▷ Your Q-learning algorithm steps go here
6: **end for**