

Machine Learning

Assignment 11.2

Submitted By: Ranji Raj

January 22, 2021

Reinforcement Learning

- Learning is only done based on the **performance feedback**.

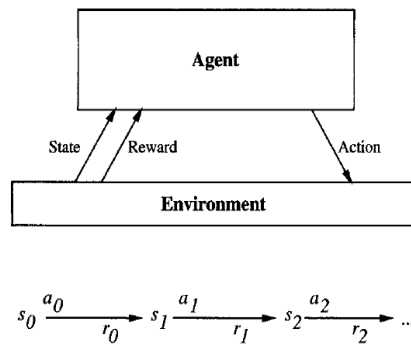


Figure 1: Learning process in RL

Learning process

- An agent observes an **environment** (or a **process**).
- At each **time step, t** the agent obtains a **state**, s_t .
- For this state the agent selects an **action**, a_t .
- At the following time step **reward**, r_{t+1} is procured.

Learning task

A control policy, $\pi : S \rightarrow A$, which maximizes the expected sum of rewards $E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots]$, with future rewards discounted exponentially by their delay.

Learning goal

Choose actions that maximize in,

$$R = \frac{1}{T} \sum_{t=1}^T r_t \quad (\text{Episodic process})$$

$$R = r_0 + \gamma r_1 + \gamma^2 r_2 + \dots = \sum_{t=1}^{\infty} \gamma^t r_t, \quad 0 \leq \gamma < 1 \quad (\text{Continuous process})$$

Comparing supervised/unsupervised/semi-supervised/RL

Supervised learning	Unsupervised learning
T: Assign labels/classes to instances	T: Assign similar instances to same cluster
E: Classify/predict output from input data	E: Learns inherent structures from data
P: Scores indicated by confusion matrix, SSE	P: Internal and external cluster validity measures
Classification: Naive Bayes, Logistic regression, Decision Tree, Perceptron, Neural Networks, kNN, LVQ Regression: Linear regression, Locally weighted regression, Regression Tree	Clustering: (<i>k</i> -Means, Fuzzy-C-Means, Hierarchical agglomerative clustering), LVQ(batch), Topic modeling, Neural Networks(autoencoders) Association rules: Apriori algorithm, Eclat, Frequent pattern-growth

Table 1: Supervised vs. Unsupervised learning

Active learning	Reinforcement learning
T: Interactive query labeling	T: Learn the optimal policy
E: Classify output from input data	E: Performance feedback
P: Scores indicated by accuracy	P: Discounted rewards
Query-by-committee, Density weighting, Variance reduction	Q-Learning, Value iteration, SARSA, A3C DQN

Table 2: Semi-Supervised vs. RL