Chapter 2
A Systematic Literature Review of Reinforcement Algorithms in Machine Learning

Gabriel Kabanda

https://orcid.org/0000-0001-6699-080X

Woxsen University, India

Hemachandran Kannan

https://orcid.org/0000-0002-7424-5022

Woxsen University, India

ABSTRACT

Reinforcement learning (RL) is learning from interactions with the environment in order to accomplish certain long-term objectives connected to the environmental condition. Reinforcement learning takes place when action sequences, observations, and rewards are used as inputs, and is hypothesis-based and goal-oriented. The purpose of the research was to conduct a systematic literature review of reinforcement algorithms in machine learning in order to develop a successful multi-agent RL algorithm that can be applied to robotics, network packet routing, energy distribution, and other applications. The robotics-related RL techniques of value-based RL, policy-based RL, model-based RL, deep reinforcement learning, meta-RL, and inverse RL were examined. As a result, the robotics-related RL techniques of value-based RL, policy-based RL, model-based RL, deep RL, meta-RL, and inverse RL were discussed in this research work. The asynchronous advantage actor-critic algorithm (A3C) is one of the best reinforcement algorithms. A3C performs better on deep RLchallenges and is quicker and easier to use.

INTRODUCTION

Machine learning is the study of statistical models and techniques that computer systems use to carry out tasks without explicit instructions (ML). By applying a number of approaches, machine learning (ML) applications teach computers how to handle data more efficiently (Hemachandran et al., 2022a). Machine learning (ML) is recognized for anticipating the strong connections between the data in large-scale data analysis. Three types of ML, according to Truong et al. (2020), include:

1. *Supervised learning* takes place when training examples are given to the algorithms as inputs labeled with the anticipated outcomes. A function is inferred using supervised learning, a machine learning task, from labeled training data made up of a selection of training instances. It entails using concrete input-output pairs to learn a function that maps an input to an output.
2. *Unsupervised learning* - When given unlabeled inputs, algorithms that participate in unsupervised learning learn on their own. Unlike supervised learning, which has right answers and a teacher, the algorithms are free to find and expose the intriguing structure in the data on their own.
3. *Reinforcement learning* occurs when action sequences, observations, and incentives are used as inputs. In order to optimize the theoretical concept of cumulative reward, reinforcement learning, a subfield of machine learning, investigates how software agents should behave in specific environments.

The process of learning from interactions with the environment with the intention of accomplishing certain long-term goals associated to the environmental situation is known as reinforcement learning (RL), and it is shown in Figure 1 with this goal in mind. Using unlabeled raw data, reinforcement learning employs an agent (system) to improve performance in response to interactions with the environment (or input attributes). The goal is set by the reward signal, which must be maximized. The reward structure is already in place before the learning process ever gets started (Hemachandran et al., 2022b). The algorithm learns patterns through trial and error, and an expert rewards it as needed. The agent must be able to fully or partially detect the environment in order to take action and alter it. So, understanding how to behave in a situation to maximize a numerical reward signal is crucial (Dadhich et al., 2021)

The main reinforcement algorithms are reviewed in this study article:

1. One-step asynchronous Q-learning
2. One-step asynchronous SARSA
3. N-step asynchronous synchronous learning
4. Actor-Critical Asynchronous Advantage (A3C)

Related Content

The Construction of Economic Intelligence as an Academic Discipline: A Theoretical Perspective
Florian Fougy (2014). *International Journal of Business Intelligence Research (pp. 1-16).*
www.irma-international.org/article/the-construction-of-economic-intelligence-as-an-academic-discipline/120049

The Role of Culture in Business Intelligence
Jore Park, Wylci Fables, Kevin R. Parkerand Philip S. Nitse (2010). *International Journal of Business Intelligence Research (pp. 1-14).*
www.irma-international.org/article/role-culture-business-intelligence/45723

Stochastic Logistics Network Design with Deadlines
Gang Wang (2014). *Encyclopedia of Business Analytics and Optimization (pp. 2277-2287).*
www.irma-international.org/chapter/stochastic-logistics-network-design-with-deadlines/107413

A Robust Multivariate Outlier Detection Method for Detection of Securities Fraud
M. Fevzi Esen (2020). *International Journal of Business Analytics (pp. 12-29).*
www.irma-international.org/article/a-robust-multivariate-outlier-detection-method-for-detection-of-securities-fraud/258268

Transforming Textual Patterns into Knowledge
Hércules Antonio do Prado, José Palazzo Moreira de Oliveira, Edilson Ferneda, Leandro Krug Wives, Edilberto Magalhaesand Stanley Loh (2004). *Business Intelligence in the Digital Economy: Opportunities, Limitations and Risks (pp. 207-227).*
www.irma-international.org/chapter/transforming-textual-patterns-into-knowledge/6072