On Selected Methods of Machine Learning for Environment

Štefan Puci, Matúš Semančík,
Manohar Gowdru Shridhara, Ľubomír Antoni,
Gabriel Semanišin

Pavol Jozef Šafárik University in Košice, Faculty of Science, Košice, Slovakia
stefan.puci@student.upjs.sk, matus.semancik@student.upjs.sk, lubomir.antoni@upjs.sk,
manohar.gowdru.shridhara@student.upjs.sk, gabriel.semanisin@upjs.sk

Machine learning methods are increasingly used to analyze complex systems in both natural and social environments. In this paper, we present an overview of selected types of machine learning (reinforcement learning, supervised learning, and unsupervised learning), and discuss their applicability to environmental monitoring and simulated social environments. Supervised learning methods are shown to support prediction and classification tasks based on labeled environmental and social data, while unsupervised learning enables the discovery of hidden structures and patterns without prior labeling. We present reinforcement learning as a framework for adaptive decision-making, allowing agents to learn optimal behavior through interaction with dynamic environments. We provide illustrative examples, including weight updates in neural networks and Q-learning updates, to clarify the learning mechanisms. The presented approaches demonstrate the practical relevance of machine learning for modeling, analyzing, and understanding socio-environmental systems.

DOI
https://doi.org/
10.18690/um.fov.3.2026.51

ISBN
978-961-299-124-1

Keywords:
machine learning,
environment,
reinforcement learning,
supervised learning,
unsupervised learning

University of Maribor Press

1 Introduction

Machine learning methods are increasingly applied in two major environmental contexts: the **natural (ecological) environment** and the **social environment**. In the **natural environment**, machine learning supports the analysis of complex environmental data such as air and water quality measurements, climate variables, satellite imagery, and biodiversity indicators. These data sources are typically high-dimensional, noisy, and heterogeneous, making traditional rule-based approaches insufficient. Machine learning models enable automated pattern detection, prediction of environmental changes, and early warning systems for pollution, extreme weather events, and ecosystem degradation (Reichstein et al., 2019; Rolnick et al., 2019).

In environmental monitoring, supervised learning is commonly used for tasks such as air pollution forecasting, land-use classification, and detection of environmental hazards from remote sensing data (Zhu et al., 2017). Unsupervised learning helps to discover hidden structures in climate and ecological datasets, for example, by identifying climate regimes, anomaly patterns, or clusters of similar ecosystems (Ghaderpour et al., 2021). Reinforcement learning has recently been explored for optimizing energy management, smart grid control, and sustainable resource allocation, where agents learn policies that minimize environmental impact while maintaining system efficiency (Vázquez-Canteli & Nagy, 2019).

In addition to natural environmental systems, machine learning is increasingly applied in the **social environment**, which encompasses interactions among autonomous agents, human-like behavior, and emergent social structures. In simulated social environments, agents are typically equipped with internal states such as goals, memory, beliefs, and social relationships. Rather than relying on explicitly scripted rules such as finite state machines or decision trees, agent behavior emerges dynamically through interaction with the environment and other agents (Wooldridge, 2009). Game-based environments provide controlled experimental settings in which such properties can be systematically studied (Yannakakis & Togelius, 2018).

2 Methods

This study is based on a literature review focusing on selected paradigms of machine learning and their applicability in natural and social environmental contexts. The search was conducted using combinations of keywords such as machine learning, environment, reinforcement learning, unsupervised learning, neural networks, agent-based simulation, and AI sustainability. Priority was given to peer-reviewed journal articles, survey papers, and influential monographs published between 2015 and 2025. Earlier foundational works (e.g., Mitchell, 1997; Sutton & Barto, 2018; Russell & Norvig, 2021) were included to provide theoretical grounding. The selection process consisted of three steps: initial screening based on title and abstract relevance, full-text evaluation of articles focusing on environmental applications, and classification of selected works according to the three primary learning paradigms: supervised learning, unsupervised learning, and reinforcement learning.

3 Machine learning in environment

Machine learning represents an area of artificial intelligence (Fig. 1) focused on the development of algorithms that enable computer systems to learn patterns from data and improve their performance without being explicitly programmed for every task (Russel and Norvig, 2021). Instead of using fixed rule-based logic, machine learning models derive relationships directly from data, making them particularly appropriate for domains such as cybersecurity and environmental monitoring.

Table 1: Main types of machine learning

	Input	Aim
Reinforcement learning	state-action pairs	Maximize future rewards across multiple time steps
Supervised learning	(**x**, y) **x:** values of input attributes y: label	Learn the function to map **x** → y
Unsupervised learning	**x** values of attributes no labels	Learn the underlying structure

Source: own table inspired by Russell & Norvig, 2021

Formally, a machine learning system learns from experience E with respect to a class of tasks T and a performance measure P, if its performance at tasks in T, as measured by P, improves with experience E (Mitchell, 1997). This definition highlights the adaptive nature of machine learning systems and their ability to generalize knowledge from observed data to unseen situations.

In recent years, the availability of large-scale datasets, increased computational power, and advances in learning algorithms have significantly expanded the applicability of machine learning. Machine learning methods are commonly categorized into three main groups according to the type of learning process they employ: **supervised learning, unsupervised learning, and reinforcement learning** (Table 1). Each of these paradigms addresses different problem settings and is suitable for distinct types of tasks in both natural and social environmental contexts (Russell & Norvig, 2021).

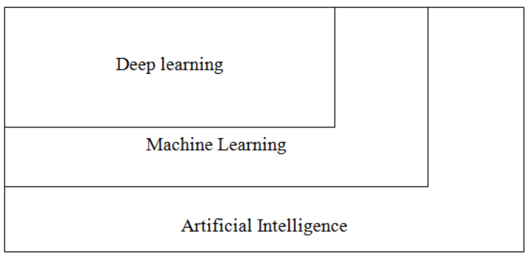

Figure 1: Position of Machine Learning in Artificial Intelligence
Source: Own diagram

3.1 Reinforcement learning

Reinforcement learning focuses on learning optimal behavior through interaction with an environment. An agent observes the current state, selects an action, and receives a reward that reflects the quality of the chosen action (Fig. 2). Over time, the agent learns a policy that maximizes cumulative reward (Sutton & Barto, 2018).

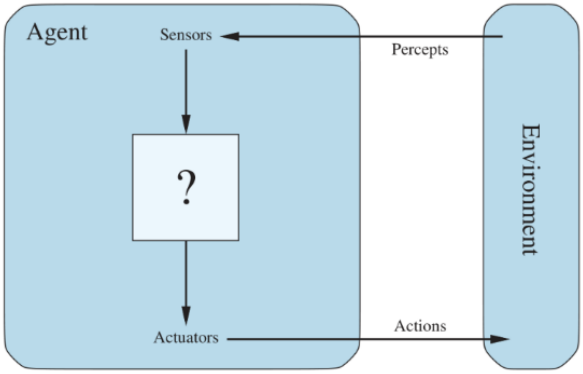

Figure 2: Agents interacting with environment
Source: Russell & Norvig, 2021

One of the most widely used reinforcement learning methods is **Q-learning**, a model-free algorithm that learns the value of taking a specific action in a given state. The learned value function, called the Q-function, estimates the expected cumulative reward for each state-action pair (Russell & Norvig, 2021).

At each time step, the agent observes the current state s, selects an action a, receives a reward r, and transitions to a new state s'. The Q-value is updated according to the following rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \quad \text{(1)}$$

where α is the learning rate, γ is the discount factor, r is the received reward, and we compute the estimated optimal future reward in the max part of equation.

In the natural environment, reinforcement learning has been applied to optimize energy management systems, control smart grids, and manage renewable energy resources. For example, a reinforcement learning agent can learn how to balance energy supply and demand in order to minimize emissions while maintaining system stability (Vázquez-Canteli & Nagy, 2019).

In the social environment, reinforcement learning can model adaptive human or agent behavior in simulated societies. Reinforcement learning agents can represent individuals or institutions that learn how to cooperate, compete, or respond to environmental policies. Such simulations help analyze how social dynamics influence sustainability outcomes and collective decision-making processes (Yannakakis & Togelius, 2018).

Consider a simulated social environment in which an agent represents an individual deciding how to behave in a group with respect to environmental cooperation. For example, the decision whether to cooperate (e.g., follow environmental regulations) or defect (e.g., ignore rules for personal benefit). Each agent observes a simplified social state, such as the level of cooperation in its neighborhood.

We can use the following states, actions, and rewards:

States (s): {low cooperation, high cooperation}

Actions (a): {cooperate, defect}

Rewards (r): +1 if cooperation improves social welfare; −1 if defection causes social harm

First, we will define our environment:

State (s = low cooperation): the current level of cooperation in the agent's neighborhood

Action (a = cooperate): The agent's behavioral choice to follow environmental rules

Reward (r = 2): positive reward for contributing to collective welfare

Next state (s' = high cooperation): The observed social state after the action.

Learning rate: $\alpha = 0.1$

The learning rate α controls the extent to which newly acquired information overrides previously learned estimates. A value of $\alpha = 0.1$ provides a balance between stability and adaptability of learning.

Discount factor: $\gamma = 0.9$

The discount factor γ determines the importance of future rewards relative to immediate ones. With $\gamma = 0.9$, the agent strongly considers long-term outcomes, which is particularly suitable for social and environmental decision-making scenarios.

Current Q-values: $\mathrm{Q}(s, a) = \mathrm{Q}(\text{low cooperation, cooperate}) = 5$

Initially, Q-values are typically initialized to zero or small random values. Through repeated interaction with the environment and exploration of available actions, the agent updates these estimates based on received rewards, gradually learning more accurate Q-values for each state–action pair.

For the next state s', the agent estimates the Q-values of available actions:

$\mathrm{Q}(s', \text{cooperate}) = 7$

$\mathrm{Q}(s', \text{defect}) = 4$

$\max\{\mathrm{Q}(s', \text{cooperate}), \mathrm{Q}(s', \text{defect})\} = 7$

<u>Now, we can update the Q-value base on Eq. (1):</u>

$\mathrm{Q}(\text{low cooperation, cooperate}) = 5 + 0.1\ [2 + 0.9 * 7 - 5]$

$\mathrm{Q}(\text{low cooperation, cooperate}) = 5.33$

After this single iteration, the estimated value of choosing cooperation in a low-cooperation social context increases from 5 to 5.33, reflecting the positive social and future rewards associated with cooperative behavior.

At the beginning of the learning process, the agent has no prior knowledge about the environment or the consequences of its actions. Therefore, Q-values are typically initialized to zero or to small random values for all state–action pairs. This neutral initialization prevents biased behavior and ensures that the agent does not prefer any action before gaining experience through interaction with the environment.

During learning, the agent explores the environment by selecting actions and observing the resulting rewards and state transitions. Exploration strategies such as the ε-greedy policy are commonly employed to balance exploration and exploitation, allowing the agent to try different actions while gradually favoring those that yield higher rewards. Each interaction provides new experience that is used to update the Q-values according to the Q-learning update rule.

Over repeated iterations, the Q-values increasingly reflect the expected long-term cumulative rewards associated with specific state–action pairs. Actions that consistently lead to favorable future outcomes are assigned higher Q-values, while less effective actions receive lower estimates. As a result, the agent progressively improves its policy and learns to select actions that maximize long-term social or environmental benefits, even in the absence of an explicit model of the environment.

Using Q-learning, each agent updates its behavior based on social feedback. When many neighbors cooperate, the effort becomes more rewarding due to social trust and shared benefits. If defection dominates, agents may adapt to protect their own interests. Agents learn policies that reflect emergent social norms, such as increased cooperation when long-term collective benefits outweigh short-term gains.

3.2 Supervised learning

Supervised learning is based on labeled training data, where each input sample is associated with a known target output (Fig. 3). The objective of the learning algorithm is to approximate a mapping function that can accurately predict outputs for unseen inputs (Mitchell, 1997). Typical supervised learning tasks include classification and regression.

In the natural environment, supervised learning is widely used for air quality prediction, climate variable forecasting, water pollution assessment, and land-cover classification from satellite imagery. For example, neural networks and support vector machines can predict concentrations of pollutants such as PM10 or NO_2 based on meteorological and emission data. Similarly, supervised image classification is applied in remote sensing to detect deforestation, urban expansion, and changes in vegetation cover (Zhu et al., 2017).

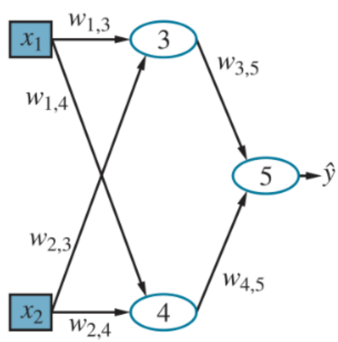

Figure 3: A neural network with two inputs and one output unit
Source: Russell & Norvig, 2021

In the social environment, supervised learning supports the analysis of human behavior, public opinion, and social responses to environmental challenges. Applications include sentiment analysis of social media posts related to climate change, prediction of pro-environmental behavior, and classification of survey responses regarding sustainability policies. These models help policymakers understand societal attitudes and design more effective environmental communication strategies (Lazer et al., 2009; Salganik, 2017).

In simulated social environments, supervised learning can be applied to model and predict agent behavior based on observed internal states and interaction histories. Consider a game-based simulation in which agents possess attributes such as goals, memory, beliefs, and social relationships, and their actions are recorded during repeated interactions (Wooldridge, 2009; Yannakakis & Togelius, 2018).

To illustrate the learning process in supervised learning, consider a simple neural network used to predict agent behavior in a simulated social environment. Let the input vector represent an agent's internal state, for example cooperation tendency and social influence, and let the target output indicate whether the agent cooperates or defects in a given situation.

Assume a single neuron with input x, weight w, and output $y' = w \cdot x$. Given a labeled training example with true target value y, the prediction error can be expressed using the squared error loss function:

$$L = 0.5(y - y')^2 \quad (2)$$

During training, the weight is updated using gradient descent to minimize the loss:

$$w \leftarrow w - \eta \frac{\partial L}{\partial w} \quad (3)$$

where η is the learning rate. For a single training step, the gradient is given by

$$\frac{\partial L}{\partial w} = -(y - y').x \quad (4)$$

and the weight update becomes

$$w \leftarrow w + \eta(y - y').x \quad (5)$$

Through repeated exposure to labeled examples, the network gradually adjusts its weights so that predicted agent behavior aligns more closely with observed outcomes. In the context of simulated social environments, this enables the model to learn how internal agent states influence social actions such as cooperation or defection, providing a transparent and interpretable mechanism for behavior prediction.

3.3 Unsupervised learning

Unsupervised learning operates on unlabeled data and aims to discover hidden structures, patterns, or relationships within the dataset. Common techniques include clustering, dimensionality reduction, and anomaly detection.

In the natural environment, unsupervised learning is used to identify climate regimes, detect environmental anomalies, and group ecosystems with similar characteristics. For instance, clustering algorithms can reveal spatial patterns in temperature or precipitation data, while anomaly detection methods can identify unusual pollution events or extreme weather phenomena. Such approaches are valuable for exploratory analysis and early warning systems (Ghaderpour et al., 2021).

In simulated social environments, unsupervised learning can be used to analyze emergent behavioral patterns arising from interactions among autonomous agents. Consider a game-based simulation in which agents are characterized by internal states such as goals, memory, beliefs, and social relationships. During the simulation, agents interact repeatedly without explicitly predefined social roles or group memberships (Wooldridge, 2009; Yannakakis & Togelius, 2018).

To illustrate unsupervised learning in a simulated social environment, consider a set of agents observed over multiple time steps in a game-based simulation. For each agent, a feature vector is constructed from behavioral statistics such as frequency of cooperation, number of interactions with other agents, and average response time. Importantly, no labels or predefined social roles are provided.

A clustering algorithm, such as *k*-means, is applied to these feature vectors in order to group agents with similar behavioral characteristics. Let x_i denote the behavioral feature vector of agent i, and let μ_k represent the centroid of cluster k. During each iteration of the algorithm, agents are assigned to the nearest cluster centroid, and the centroids are updated as the mean of all agents assigned to the cluster C_k:

$$\mu_k \leftarrow \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i \quad \text{……………………………………………………(6)}$$

Through repeated iterations, the clustering process converges to a partition of the agent population into behaviorally coherent groups. These clusters may correspond to emergent social roles such as cooperative agents, dominant influencers, or socially isolated individuals, despite the absence of any explicit role definitions.

4 Discussion and conclusion

The comparison of machine learning paradigms highlights distinct advantages and limitations in environmental contexts. Across all paradigms, common challenges include: data scarcity and imbalance, limited interpretability of deep learning models, sensitivity to environmental non-stationarity, transferability issues across geographic regions (e.g., Mitchell, 1997; Sutton & Barto, 2018; Russell & Norvig, 2021).

This paper presented an overview of selected machine learning methods and their applicability in both natural and social environmental contexts. Through illustrative examples, the paper demonstrated how supervised learning enables prediction and classification tasks based on labeled data, how unsupervised learning supports the discovery of hidden structures and emergent patterns, and how reinforcement learning allows agents to adapt their behavior through interaction with dynamic environments. Special attention was given to simulated social environments, where machine learning methods provide a data-driven framework for modeling agent behavior and analyzing social dynamics without relying on predefined rules.

Acknowledgment

This work was supported by the VVGS ESGD grant (Early Stage Grants – Pavol Jozef Šafárik University in Košice), led by Ing. Manohar Gowdru Shridhara, ŠPP element 0HV040126 / FS 190180 / 2887 – VVGS ESGD, code 09I03-03-V05-00008. (M. Gowdru Shridhara). This work was supported by the the project vvgs-2025-3506 Dynamics of Multi-Agent Behavior in Simulated Environments using Reinforcement Learning given by Internal Research Grant System of the Faculty of Science, Pavol Jozef Šafárik University in Košice, within the call VVGS PF UPJŠ 2025-2026 (Š. Puci, M. Semančík). This work was supported by the Scientific Grant Agency of the Ministry of Education, Research, Development and Youth of the Slovak Republic and the Slovak Academy of Sciences under the project VEGA 1/0780/24 "Combination of LiDAR and hyperspectral data with machine learning methods for improving land cover classification."(L. Antoni)

References

Epstein, J. M. (2006). Generative social science: Studies in agent-based computational modeling. Princeton University Press.

Ghaderpour, E., Pagiatakis, S. D., & Hassan, Q. K. (2021). A survey on change detection and time series analysis with applications. Applied Sciences, 11(13), 6141. doi:10.3390/app11136141

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Van Alstyne, M. (2009). Computational social science. Science, 323(5915), 721–723. doi:10.1126/science.1167742

Mitchell, T. M. (1997). Machine learning. McGraw-Hill.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566, 195–204. doi:10.1038/s41586-019-0912-1

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., et al. (2019). Tackling climate change with machine learning. arXiv preprint, arXiv:1906.05433.

Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson Education.

Salganik, M. J. (2017). Bit by bit: Social research in the digital age. Princeton University Press.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied Energy, 235, 1072–1089. doi:10.1016/j.apenergy.2018.11.002

Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). Wiley.

Yannakakis, G. N., & Togelius, J. (2018). Artificial intelligence and games. Springer.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36. doi:10.1109/MGRS.2017.2762307

About the authors

Štefan Puci is a PhD student at Faculty of Science, Pavol Jozef Šafárik University in Košice. His research interests focus on reinforcement learning, AI agents, and other methods of machine learning.

Matúš Semančík is a PhD student at the Faculty of Science, Pavol Jozef Šafárik University in Košice. His research interests include dynamical systems and fuzzy logic.

Manohar Gowdru Shridhara is a PhD student at the Faculty of Science, Pavol Jozef Šafárik University in Košice. His research interests include machine learning and optimization techniques.

Ľubomír Antoni is an associate professor at the Institute of Computer Science, Faculty of Science, Pavol Jozef Šafárik University in Košice. His research interests include artificial intelligence, fuzzy systems, data mining, and applied machine learning.

Gabriel Semanišin is a professor of Computer Science at Faculty of Science, Pavol Jozef Šafárik University in Košice. As part of his research activities, he focuses mainly on algorithmic graph theory and its application in various areas of theoretical and applied informatics. He is a co-guarantor of the study programs Applied Informatics, Data Analysis and Artificial Intelligence and Computer Science. He was a supervisor of 6 PhD students in study programs Computer Science, Discrete Mathematics and Theory of Teaching Informatics.