What is AI agent learning?

How do AI agents learn and adapt over time?

AI agent learning refers to the process by which an artificial intelligence (AI) agent improves its performance over time by interacting with its environment, processing data and optimizing its decision-making. This learning process enables autonomous agents to adapt, improve efficiency and handle complex tasks in dynamic environments. Learning is a fundamental component of many agentic AI systems.

Not all AI agent types can learn. Some are simple reflex agents that passively take in data and lacking learning capabilities, perform reactive programmed actions in response.

There are model-based reflex agents that can reason about their environment, and proactive goal-based agents that can pursue specific goals, but they do not learn. Neither can utility-based agents, which use a utility function to evaluate and select actions that maximize overall benefit.

Industry newsletter

The latest AI trends, brought to you by experts

Get curated insights on the most important—and intriguing—AI news. Subscribe to our weekly Think newsletter. See the IBM Privacy Statement.
We use your email to validate you are who you say you are, to create your IBMid, and to contact you for account related matters.
Business email

Your subscription will be delivered in English. You will find an unsubscribe link in every newsletter. You can manage your subscriptions or unsubscribe here. Refer to our IBM Privacy Statement for more information.

Your subscription will be delivered in English. You will find an unsubscribe link in every newsletter. You can manage your subscriptions or unsubscribe here. Refer to our IBM Privacy Statement for more information.

What is a learning agent?

A learning agent improves its performance over time by adapting to new experiences and data. Other AI agents work with predefined rules or models, while learning agents continuously update their behavior based on feedback from the environment.

This allows them to enhance their decision-making abilities and perform better in dynamic and uncertain situations. Learning agents represent the full potential of AI tools to handle multistep problem-solving workloads with minimal human intervention.

Learning agents typically consist of six key components:

1. **Sensors or perceptors:** Perceive the agent's environment and send data to the agent to enable decision-making and behavioral adjustments.

2. **Actuators:** The means through which the agent interacts with its environment, such as API integrations.

3. **Performance element**: Makes informed decisions based on a knowledge base.

4. **Learning element**: Adjusts and improves the agent's knowledge based on feedback and experience.

5. **Critic**: Evaluates the agent's actions and provides feedback, often in the form of rewards or penalties.

6. **Problem generator**: Suggests exploratory actions to help the agent discover new strategies and improve its learning.

Types of AI agent learning

Machine learning (ML) forms the backbone of the various types of AI agent learning. It enables agents to identify patterns, make predictions and improve performance based on data.

The three primary machine learning techniques used in AI agents are supervised learning, unsupervised learning and reinforcement learning. These are deep learning techniques that use complex neural networks with many layers to process vast amounts of data and learn intricate patterns.

Supervised learning

Supervised learning involves training machine learning algorithms on labeled datasets, where each input corresponds to a known output. The agent uses this information to build predictive models.

For example, AI chatbots can be trained on customer service conversations and corresponding resolutions to provide predicted responses. This approach is widely applied in image recognition, speech-to-text processing and medical diagnostics.

Transfer learning allows AI agents to use knowledge acquired from one task and apply it to another. For instance, a large language model (LLM) trained on a general dataset can be fine-tuned for a specific domain, such as legal or medical text processing.

Unsupervised learning

In contrast, unsupervised learning allows AI agents to perform data analysis on unlabeled data to find patterns and structures without human oversight.

This method is useful in tasks such as clustering customer behavior to improve marketing strategies,

anomaly detection in cybersecurity and recommendation systems such as those used by streaming services.

Self-supervised learning uses unsupervised learning for tasks that conventionally require supervised learning. Rather than relying on labeled datasets for supervisory signals, self-supervised AI models generate implicit labels from unstructured data.

Self-supervised learning is useful in fields such as computer vision and natural language processing(NLP), which require large amounts of labeled training data.

Reinforcement learning

Reinforcement learning (RL) is a machine learning process that focuses on decision-making workflows in autonomous agents. It addresses sequential decision-making processes in uncertain environments.

In contrast to supervised learning, reinforcement learning does not use labeled examples of correct or incorrect behavior. However, reinforcement learning also differs from unsupervised learning in that reinforcement learning learns by trial-and-error and reward function rather than by extracting information of hidden patterns.

Reinforcement learning is also distinct from self-supervised learning because it does not produce pseudo labels or measure against a ground truth—it is not a classification method but an action learner.

AI agents using reinforcement learning operate through a trial-and-error process, where they take actions within an environment, observe the outcomes and adjust their strategies accordingly. The learning process involves defining a policy that maps states to actions, optimizing for long-term cumulative rewards rather than immediate gains.

Over time, the agent refines its decision-making capabilities through repeated interactions, gradually improving its ability to perform complex tasks effectively. This approach is beneficial in dynamic environments where predefined rules might not be sufficient for optimal performance.

Autonomous vehicles use reinforcement learning to learn optimal driving behaviors. Through trial and error, the AI improves its ability to navigate roads, avoid obstacles and make real-time driving decisions. AI-powered chatbots improve their conversational abilities by learning from user interactions and optimizing responses to enhance engagement.

Continuous learning

Continuous learning in AI agents refers to the ability of an artificial intelligence system to learn and adapt over time, incorporating new data and experiences without forgetting previous knowledge. For example, a game-playing agent could improve its skills with each game, or even with each move.

Unlike traditional machine learning, which typically involves training on a fixed dataset, continuous learning enables the AI to update its models continuously as it encounters new information or changes in its environment. An adaptive agent can improve its performance in real time, adapting to new patterns, changing environments, evolving situations and dynamic conditions.

Continuous learning is important in real-world applications where data is constantly changing and the AI must stay up to date with new inputs to remain effective. It helps prevent "catastrophic forgetting," where the model forgets old knowledge when learning new information and helps ensure that the system can handle an ever-evolving set of tasks and challenges.

Multiagent learning and collaboration

One of the benefits of AI agents is that they can work together. In multiagent architectures, AI agents learn through collaboration and competition. In cooperative learning, agents share knowledge to achieve a common goal, as seen in swarm robotics.

However, competitive learning occurs when agents refine their strategies by competing in adversarial settings, such as financial trading AI.

Imagine a network of AI agents working to improve patient care, streamline workflows, promote adherence to ethical considerations and optimize resource allocation in a hospital network.

In multiagent systems, sometimes a more advanced learning agent equipped with generative AI (gen AI) oversees simpler reflexive or goal-based agents. The simpler agents might accomplish repetitive tasks assigned by the learning agent. In this use case, each agent could represent a different role or task within the healthcare system, and they would collaborate and share information to enhance patient outcomes and operational efficiency.

Feedback mechanisms

With feedback mechanisms, an AI system receives information about the results of its actions or predictions, allowing it to assess the accuracy or effectiveness of its behavior.

This feedback, which can be positive (reinforcing correct behavior) or negative (penalizing incorrect behavior), is essential for guiding the system's decisions and improving its performance. Feedback is a critical component that enables learning in AI, but it is not the entirety of the learning process.

Real-time feedback is crucial for AI agents operating in dynamic environments. Autonomous systems, such as self-driving cars and robotic process automation (RPA), continuously gather sensor data and adjust their behavior based on immediate feedback. This allows them to adapt to changing conditions and improve their real-time decision-making.

Unsupervised learning feedback

In unsupervised learning, feedback is not explicitly provided in the form of labeled data or direct supervision. Instead, the AI agent seeks patterns, structures or relationships within the data itself.

For instance, in clustering or dimensionality reduction tasks, feedback occurs implicitly as the agent adjusts its model to best represent the underlying structure of the data.

The model refines its understanding of the data through metrics such as error minimization, for example, reducing reconstruction error in autoencoders or optimizing a specific criterion, such as maximizing data similarity in clustering.

In a supply chain management system that needs to predict product demand and optimize inventory levels across multiple warehouses and stores, an AI agent could use unsupervised learning techniques, such as clustering or anomaly detection to analyze large volumes of historical sales data, without the need for explicit labels or predefined categories.

Supervised learning feedback

In supervised learning, feedback is explicit and comes in the form of labeled data. The AI agent is trained using input/output pairs (for example, an image with a corresponding label). After the agent makes predictions, feedback is provided by comparing its output to the correct label (ground truth).

The difference between the predicted and true output (error) is calculated, often using a loss function. This feedback is then used to adjust the model parameters so that the model can improve its predictions over time.

AI agents can use supervised learning to predict which products or services a customer is likely to be interested in, based on their past behavior, purchase history or user preferences.

For example, if building AI agents for e-commerce apps, historical data such as past purchases and ratings as labeled examples could be used to to train a model that predicts the products a customer might want to purchase next, improving customer experiences.

Supervised learning is considered human-in-the-loop (HITL) learning because AI agents integrate human feedback to refine their models, improve decision-making and adapt to new situations.

This method combines automated learning with human expertise, allowing AI to handle complex tasks more effectively while minimizing errors and biases. HITL can also be integrated as a feedback mechanism in other types of learning, but it is only integral to the process of self-supervised learning.

Reinforcement learning feedback

In reinforcement learning, feedback is provided in the form of rewards or penalties. An RL agent interacts with an environment, performing actions that lead to different outcomes. After each action, the agent receives feedback in the form of a scalar reward or penalty that indicates how good or bad the outcome was relative to the goal.

The agent uses this feedback to adjust its policy or decision-making strategy, aiming to maximize cumulative rewards over time. This feedback loop allows the agent to learn optimal actions or strategies through trial and error, refining its behavior as it explores the environment.

Self-supervised learning feedback

In self-supervised learning, the agent generates its own labels from the data, creating a form of feedback from the structure within the data itself. The model uses parts of the data to predict other parts, such as predicting missing words in a sentence or predicting future frames in a video.

The feedback comes from comparing the model's predictions to the actual missing or future data. The agent learns by minimizing the prediction error, refining its internal representations based on this self-generated feedback.