Lesson 4

Lesson 4: Machine Learning

LESSON SKILLS

After completing this lesson, you will be able to:

- Explain how machine learning uses statistical learning algorithms to build smart systems.
- Describe the five stages of machine learning.
- Compare and contrast supervised and unsupervised learning.
- Describe reinforcement learning.
- Explain how data is the core of machine learning algorithms.
- Identify the three functions of a machine learning system.
- Understand the difference between classification, regression and clustering algorithms.

KEY TERMS

- Big Data
- Classification Algorithms
- Deep Learning
- Machine Learning
- Regression Algorithms
- Reinforcement Learning
- Supervised Learning
- Unsupervised Learning

Points to Ponder

These Points to Ponder are designed to help you focus on key elements in this lesson. They are also suitable for use to spark discussions or individual research.

- How does machine learning use statistical learning algorithms to build smart systems?
- Recommendation systems can suggest songs you may like. How does that happen?
- Why do software applications become more accurate in predicting with machine learning?
- Describe the four stages of machine learning. Compare and contrast supervised and unsupervised learning.
- Describe reinforcement learning.
- Explain how data is the core of machine learning algorithms.
- How does machine learning relate to data science?

Overview

This lesson introduces machine learning. You will learn that AI uses data to work, and that machine learning is a subset of AI. Machine learning can be supervised and unsupervised, and you will learn the difference between the two and what reinforcement learning is. There are four stages of machine learning and three functions of machine learning. Machine learning has many benefits and can be so helpful in today's world. This lesson shows some examples of how machine learning is making a difference in the quality of life and the betterment of the planet.

Defining Machine Learning

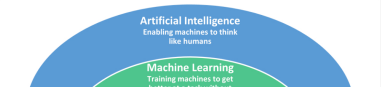
Objectives
5.1: How does machine learning fit into AI
5.1.1: Define machine learning
5.1.2: Describe how artificial intelligence applies machine learning
5.1.4: Explain how data collection is the first step in machine learning
5.1.7: Identify the three functions of a machine learning system (descriptive, predictive and prescriptive)

Think About This

At the core of machine learning is data. Why is data so easily collected and manipulated in today's world? How does the development of the Internet and computer technology allow machine learning to happen?

There are different definitions of machine learning, some of which are more complex than others. Simply put, it is the field of study that gives computers the ability to learn without being explicitly programmed.

Remember it is a subset of AI. While AI is the broad science of mimicking human abilities attempting to have machines “think” like humans, machine learning is a specific subset of AI that trains a machine how to learn.



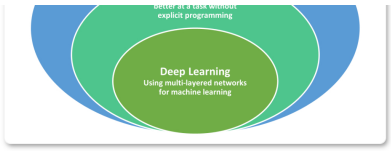


Figure 4-1: Subsets of AI

Machine learning is an application of artificial intelligence (AI) that gives systems the ability to “automatically” learn and get better from experience without being specifically programmed. Machine learning concentrates on the development of computer programs that can access data and use it to learn for themselves. Machine learning is very important because it allows the use of algorithms and data to have the computer system “learn”. It is a branch of artificial intelligence based on the notion that systems can learn from data and also identify patterns and make decisions with very little human intervention.

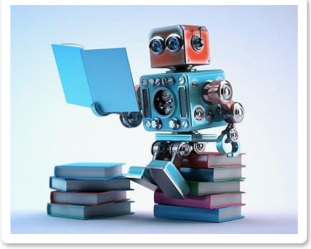


Figure 4-2: Machine learning

Machine learning starts with data. Data can be numbers, text, images, (like pictures of people), transaction purchase numbers, sales reports, or information from cameras or sensors. The data is collected and turned into training data. The more data, the better the program. Machine learning relies on large amounts of data, big data! Big data is a collection of large and complex data sets. Big data is collected from a variety of sources like social media, voice searches, transactions, satellite imagery, wearable devices, the Internet of things, email tracking, and online webpages. Big data consists of three things often called the three V’s. These are the defining properties of big data.

- **Volume** – the amount of data
- **Velocity** – how fast the data comes in
- **Variety** – different forms of data- (text, numbers, images, video)

Machine learning improves itself without being explicitly programmed. It uses the growing and changing input of data to reach its goals and serve the purpose it was designed for.

Figure 4-3: Data input to improve its purpose

Machine learning is using data to solve problems and answer questions. Machine learning has three different “functions” or purposes. One purpose is that the machine learning system can be descriptive. This means the system uses data to explain what happened. The next function is predictive where the machine learning system uses the data to predict what will happen. The last one is prescriptive which means the system will use the data to make recommendations or suggestions about what to do next. In the table below, are examples of each function and how they are helpful.

Type of Function/ Purpose	Example	How it is Helpful
Descriptive	Sales report	Helps a store see how much of specific products were sold
Predictive	Credit score report	Based on your past financial behavior, it can estimate your ability to be eligible for a loan
Prescriptive	Customer experience survey	Helps companies recommend specific customer experiences so they will be the most satisfied

Table 4-1: Functions of Machine Learning

With the help of Machine Learning, we can develop intelligent systems that are capable of taking decisions on an autonomous basis. These algorithms learn from the past instances of data through statistical analysis and pattern matching. Then, based on the learned data, it provides us with the predicted results.

Data is the core backbone of machine learning algorithms. With the help of the historical data, we are able to create more data by training these machine learning algorithms.

For example, Generative Adversarial Networks (GANs) are an advanced concept of machine learning that learns from the historical images through which they are capable of generating more images. This is also applied towards speech and text synthesis.

Therefore, machine learning has opened up a vast potential for data science applications. Machine learning combines computer science, mathematics, and statistics. Statistics is essential for drawing inferences from the data.

In figure 4-4, data is input then classified by a human into machine learning and the output results are checked. With deep learning the data is extracted automatically, and the machine learns from its errors without human involvement.

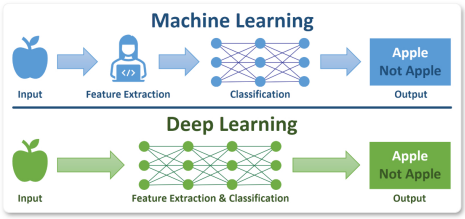


Figure 4-4: Machine learning vs. deep learning

Link to Learn More

Read more about machine learning and big data

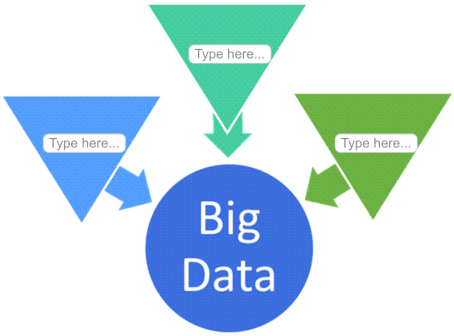
- Machine learning facts
- Big data facts
- What is Machine Learning? (YouTube 2 minutes)

Suggested activities

- The Three V's of Big Data (See below)
- Teachable Machine (Online)
- Where Does Big Data Come From? (Hands-on)
- Functions of Machine Learning (See below)

The Three V's of Big Data — FULL SCREEN RESET SUBMIT

Type each of the Three V's in the fields below.



Functions of Machine Learning — FULL SCREEN RESET SUBMIT

Fill in the blanks with the options on the dropdown lists.

One purpose is that the machine learning system can be Please Select. This means the system uses data to explain what happened.

The next function is Please Select where the machine learning system uses the data to predict what will happen.

The last one is Please Select which means the system will use the data to make recommendations or suggestions about what to do next.

Demonstration videos

- What is Machine Learning

Types of Machine Learning

Objectives

5.1.5: Identify examples of machine learning
5.1.6: Explain how machine learning works
5.2.1: Define supervised learning
5.2.2: Define unsupervised learning
5.2.3: Define reinforcement learning
5.2.4 Describe how machines use data differently in each category of machine learning
5.2.5: Describe the difference between classification and regression algorithms in supervised learning

TABLE OF CONTENTS

5.2.6: Define clustering algorithms in unsupervised learning

We know that machine learning is an application of AI where machines access data and then use that data to learn for themselves. It's simply put, getting a computer to perform a task without explicitly being programmed to do so. Instead of coding software with specific instructions, machine learning trains an algorithm so it can learn how to make decisions for itself. Machine learning explores the study and creation of algorithms that can learn and make predictions on data. These algorithms follow programmed instructions but can also make predictions or decisions based on data.

In order to train an algorithm, you need a neural network, which is a set of algorithms inspired by and modeled after the human brain, which is made up of individual neurons connected to each other. In machine learning, a **neuron** is an interconnected element that processes inputs. A neuron receives data, processes it and then sends the result.

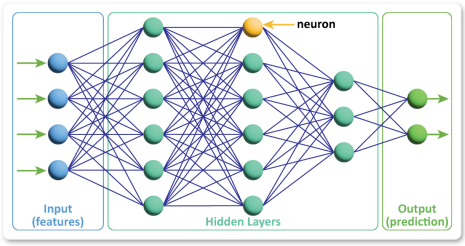


Figure 4-5: Neural network

Machine learning uses these neurons for a variety of tasks like predicting the outcomes of dramatic weather for example, hurricanes and floods. It can even predict the movement of a basketball player during a game. A neuron uses input data from any past events to predict the outcome. There are three categories of machine learning: supervised learning, unsupervised and reinforcement learning.

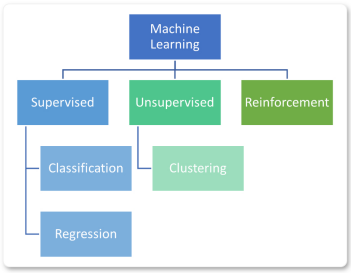


Figure 4-6: Types of machine learning

Supervised Learning

One of the main categories of machine learning problems is supervised learning. Supervised learning is one of the most basic types of machine learning. In this type, the machine learning algorithm is trained on labeled data. This is where there is available training data so the program can get feedback on its performance as it learns. Things like playing games and identifying objects are considered supervised learning because the computer is getting feedback as it learns. Did it guess what the object was correctly? Did it get a good score in the game, or lose quickly? Feedback allows it to adjust its decision-making process so it can do better next time. Because it is getting feedback, it is categorized as “supervised” learning.

One of the most common sub-categories of supervised learning problems is classification. In a classification problem, the program is given a set of inputs and must learn to classify those inputs correctly, like an email spam filter or image recognition program. An example of this could be teaching a machine to recognize a picture of a dog. You would train the machine by showing it pictures of various breeds of dogs, labeled as dogs as compared to pictures of cats, labeled as cats. When it comes across a picture of a dog, it would recognize it as a dog based on the data on which it had been trained.

Regression algorithms are used to predict the continuous values such as price, salary, or age. Regression algorithms can be used to solve the regression problems such as weather prediction, or house price prediction. These are used with continuous data like wind speed or temperature, where classification ones are used with discrete data, which means there is countable, measurable distinct data like number of likes on Facebook or how many votes for a person.

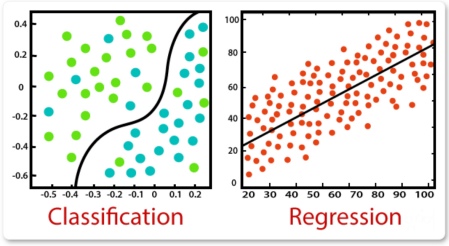


Figure 4-7: Sub-categories of supervised learning

Unsupervised Learning

In unsupervised learning , the machine is not provided labeled examples or previous patterns on which to base the analysis of the data inputs. The machine must uncover patterns and draw inferences by itself without having the correct answers. It will classify or cluster data by discovering the similarity of features on its own. Using unsupervised learning, the machine would be fed millions of pictures of dogs, without labeling them as dogs. It would use the text or captions associated with the pictures to decipher clues, particularly attending to the fact that the word dog often showed up in the various texts and would therefore (on its own) label the photos as dogs.

Clustering is an unsupervised machine learning method of identifying and grouping similar data points in larger datasets without concern for the specific outcome. Clustering or cluster analysis is a machine learning technique, which groups the unlabeled dataset

It does it by finding some similar patterns in the unlabeled dataset such as shape, size or color.

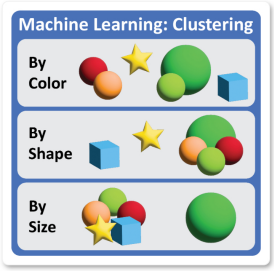


Figure 4-8: Machine learning clustering

Reinforcement Learning

Reinforcement learning is different from both supervised and unsupervised learning because it continuously improves based on feedback from experiences. It learns through trial and error – from the consequences of its action and by the results of making new choices. As an action is taken, the success of the outcome is graded and receives either a positive or negative score. The algorithm seeks to receive positive scores and the model is trained on continuous feedback. An example of this could be a self-driving car for which getting from one location to another without crashing would receive a positive score. It would have to also follow the rules of the road to get positive scores like staying within the speed limit, changing lanes successfully, and obeying traffic lights.

Here are some other reinforcement learning examples:

- Robots completing tasks - stocking warehouse shelves
- Running ads - matching ads with consumers
- Financial trading - stock market trading
- Marketing strategies - uses reactions of customers
- Online recommendations - recommending videos
- Industrial machine teaching - independent forklifts

Figure 4-9: Independent forklift

Machine Learning Training Stages

Objectives
5.1.3: Identify the five stages of machine learning training
5.1.6: Explain how machine learning works
8.1.4 Describe how training data affects the accuracy of supervised machine learning

Machine learning is used to find solutions to various challenges that arise across a variety of different scenarios and environments. There are five stages of machine learning training. The first stage is when data gets put in. The next is where it is analyzed then patterns are found and finally results happen. Depending on the type of machine learning the results of the final stage vary. The five stages of training are:

1. **Get Data** – collect the data needed.
2. **Prepare Data** – clean the data and remove any unwanted.
3. **Train Model** – pass the prepared data to your machine learning model to find patterns and make predictions. This is the most important step in machine learning.
4. **Test Data** – see if it gives the desired results.
5. **Improve** – make adjustments based on the results.

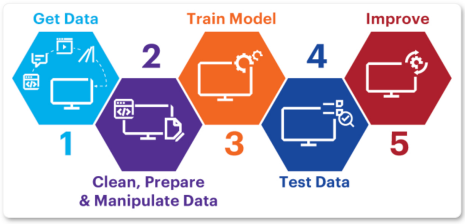


Figure 4-10: Five stages of machine training

Suggested activity:

- Machine Learning Training Stages (See below)

Machine Learning Training Stages FULL SCREEN RESET SUBMIT

Put the stages of machine learning training in order.

Machine Learning Training Stages

- Train Model
- Test Data
- Prepare Data
- Improve
- Get Data

Click to select. Drag and Drop to set sequence.

Applications of Machine Learning

Objective
5.1.5: Identify examples of machine learning

Machine learning is used extensively and in many implementations of AI. Because of its ability to analyze massive amounts of data, provide information, and make meaningful predictions and recommendations, it enables people to use it for great purposes.

Machine learning is all around us and it has been used for quite some time. Some famous people have started their computer science work in machine learning projects. During Mark Zuckerberg's high school years, he worked under the company name Intelligent Media Group to build a music player called the Synapse Media Player. The device used machine learning to learn the user's listening habits.

Figure 4-11: Synapse Media Player

In the late 1950s, Arthur Samuel coined the phrase "machine learning" and wrote the first computer learning program, which was designed for the game of checkers (the IBM computer improved at the game the more it played). Since the beginning of machine learning until the present day it is used to solve problems.

Figure 4-12: Applications of machine learning

Machine learning as a career

Machine learning has practical business applications like analyzing large volumes of data, powering self-driving vehicles, and assisting medical diagnoses. As AI research advances, the number of tasks it can perform will only increase. Companies are already in great need for AI experts and are hiring those with expertise in the field.

Link to Learn More

Read more about Mark Zuckerberg and Arthur Samuel and review the top applications used today.

- Mark Zuckerberg facts
- Arthur Samuel facts
- Top 5 Applications of Machine Learning (Slideshow)
- Top 10 Applications of Machine Learning (YouTube 5 minutes)

Suggested activities

- Hour of Code Activities (Online)
- Research a Machine Learning Application (Hands-on)

Demonstration videos

- AI is Here Now
- Top 10 Applications of Machine Learning

Glossary

Classification Algorithms

Machine learning methods used to categorize data into predefined groups or classes based on patterns in the data.

Deep learning

A type of machine learning that uses layered artificial neural networks to automatically learn patterns and make decisions from large amounts of data.

Machine learning

A type of AI that enables computers to learn from data and improve their performance on tasks without being explicitly programmed.

Reinforcement Learning

A type of machine learning where an agent learns by interacting with an environment, receiving rewards or penalties, and improving its actions to maximize long-term rewards.

Supervised Learning