IN THIS CHAPTER

» Defining AI and its history

» Using AI for practical tasks

» Seeing through AI hype

» Connecting AI with computer technology

Chapter 1
Introducing AI

Artificial Intelligence (AI) has had several false starts and stops over the years, partly because people don't really understand what AI is all about, or even what it should accomplish. A major part of the problem is that movies, television shows, and books have all conspired to give false hopes as to what AI will accomplish. In addition, the human tendency to *anthropomorphize* (give human characteristics to) technology makes it seem as if AI must do more than it can hope to accomplish. So, the best way to start this book is to define what AI actually is, what it isn't, and how it connects to computers today.

REMEMBER

Of course, the basis for what you expect from AI is a combination of how you define AI, the technology you have for implementing AI, and the goals you have for AI. Consequently, everyone sees AI differently. This book takes a middle-of-the-road approach by viewing AI from as many different perspectives as possible. It doesn't buy into the hype offered by proponents, nor does it indulge in the negativity espoused by detractors. Instead, it strives to give you the best possible view of AI as a technology. As a result, you may find that you have somewhat different expectations than those you encounter in this book, which is fine, but it's essential to consider what the technology can actually do for you, rather than expect something it can't.

Defining the Term AI

Before you can use a term in any meaningful and useful way, you must have a definition for it. After all, if nobody agrees on a meaning, the term has none; it's just a collection of characters. Defining the idiom (a term whose meaning isn't clear from the meanings of its constituent elements) is especially important with technical terms that have received more than a little press coverage at various times and in various ways.

REMEMBER

Saying that AI is an artificial intelligence doesn't really tell you anything meaningful, which is why there are so many discussions and disagreements over this term. Yes, you can argue that what occurs is artificial, not having come from a natural source. However, the intelligence part is, at best, ambiguous. Even if you don't necessarily agree with the definition of AI as it appears in the sections that follow, this book uses AI according to that definition, and knowing it will help you follow the rest of the text more easily.

Discerning intelligence

People define intelligence in many different ways. However, you can say that intelligence involves certain mental activities composed of the following activities:

- **Learning:** Having the ability to obtain and process new information
- **Reasoning:** Being able to manipulate information in various ways
- **Understanding:** Considering the result of information manipulation
- **Grasping truths:** Determining the validity of the manipulated information
- **Seeing relationships:** Divining how validated data interacts with other data
- **Considering meanings:** Applying truths to particular situations in a manner consistent with their relationship
- **Separating fact from belief:** Determining whether the data is adequately supported by provable sources that can be demonstrated to be consistently valid

The list could easily get quite long, but even this list is relatively prone to interpretation by anyone who accepts it as viable. As you can see from the list, however, intelligence often follows a process that a computer system can mimic as part of a simulation:

1. Set a goal based on needs or wants.
2. Assess the value of any currently known information in support of the goal.
3. Gather additional information that could support the goal. The emphasis here is on information that could support the goal, rather than information that you know will support the goal.
4. Manipulate the data such that it achieves a form consistent with existing information.
5. Define the relationships and truth values between existing and new information.
6. Determine whether the goal is achieved.
7. Modify the goal in light of the new data and its effect on the probability of success.
8. Repeat Steps 2 through 7 as needed until the goal is achieved (found true) or the possibilities for achieving it are exhausted (found false).

REMEMBER

Even though you can create algorithms and provide access to data in support of this process within a computer, a computer's capability to achieve intelligence is severely limited. For example, a computer is incapable of understanding anything because it relies on machine processes to manipulate data using pure math in a strictly mechanical fashion. Likewise, computers can't easily separate truth from mistruth (as described in Chapter 2). In fact, no computer can fully implement any of the mental activities described in the list that describes intelligence.

As part of deciding what intelligence actually involves, categorizing intelligence is also helpful. Humans don't use just one type of intelligence, but rather rely on multiple intelligences to perform tasks. Howard Gardner of Harvard has defined a number of these types of intelligence (see the article "Multiple Intelligences" from Project Zero at Harvard University for details), and knowing them helps you to relate them to the kinds of tasks that a computer can simulate as intelligence (see Table 1-1 for a modified version of these intelligences with additional description).

Discovering four ways to define AI

As described in the previous section, the first concept that's important to understand is that AI doesn't really have anything to do with human intelligence. Yes, some AI is modeled to simulate human intelligence, but that's what it is: a simulation. When thinking about AI, notice an interplay between goal seeking, data

TABLE 1-1 **The Kinds of Human Intelligence and How AIs Simulate Them**

Type	Simulation Potential	Human Tools	Description
Visual-spatial	Moderate	Models, graphics, charts, photographs, drawings, 3-D modeling, video, television, and multimedia	Physical-environment intelligence used by people like sailors and architects (among many others). To move at all, humans need to understand their physical environment — that is, its dimensions and characteristics. Every robot or portable computer intelligence requires this capability, but the capability is often difficult to simulate (as with self-driving cars) or less than accurate (as with vacuums that rely as much on bumping as they do on moving intelligently).
Bodily-kinesthetic	Moderate to High	Specialized equipment and real objects	Body movements, such as those used by a surgeon or a dancer, require precision and body awareness. Robots commonly use this kind of intelligence to perform repetitive tasks, often with higher precision than humans, but sometimes with less grace. It's essential to differentiate between human augmentation, such as a surgical device that provides a surgeon with enhanced physical ability, and true independent movement. The former is simply a demonstration of mathematical ability in that it depends on the surgeon for input.
Creative	None	Artistic output, new patterns of thought, inventions, new kinds of musical composition	Creativity is the act of developing a new pattern of thought that results in unique output in the form of art, music, and writing. A truly new kind of product is the result of creativity. An AI can simulate existing patterns of thought and even combine them to create what appears to be a unique presentation but is really just a mathematically based version of an existing pattern. In order to create, an AI would need to possess self-awareness, which would require intrapersonal intelligence.
Interpersonal	Low to Moderate	Telephone, audio conferencing, video conferencing, writing, computer conferencing, email	Interacting with others occurs at several levels. The goal of this form of intelligence is to obtain, exchange, give, and manipulate information based on the experiences of others. Computers can answer basic questions because of keyword input, not because they understand the question. The intelligence occurs while obtaining information, locating suitable keywords, and then giving information based on those keywords. Cross-referencing terms in a lookup table and then acting on the instructions provided by the table demonstrates logical intelligence, not interpersonal intelligence.

Type	Simulation Potential	Human Tools	Description
Intrapersonal	None	Books, creative materials, diaries, privacy, and time	Looking inward to understand one's own interests and then setting goals based on those interests is currently a human-only kind of intelligence. As machines, computers have no desires, interests, wants, or creative abilities. An AI processes numeric input using a set of algorithms and provides an output; it isn't aware of anything that it does, nor does it understand anything that it does.
Linguistic (often divided into oral, aural, and written)	Low for oral and aural None for written	Games, multimedia, books, voice recorders, and spoken words	Working with words is an essential tool for communication because spoken and written information exchange is far faster than any other form. This form of intelligence includes understanding oral, aural, and written input, managing the input to develop an answer, and providing an understandable answer as output. In many cases, computers can barely parse input into keywords, can't actually understand the request at all, and output responses that may not be understandable at all. In humans, oral, aural, and written linguistic intelligence come from different areas of the brain (see "**Say What? How the Brain Separates Our Ability to Talk and Write**" from John Hopkins University), which means that even with humans, someone who has high written linguistic intelligence may not have similarly high oral linguistic intelligence. Computers don't currently separate aural and oral linguistic ability — one is simply input and the other output. A computer can't simulate written linguistic capability because this ability requires creativity.
Logical-mathematical	High (potentially higher than humans)	Logic games, investigations, mysteries, and brain teasers	Calculating a result, performing comparisons, exploring patterns, and considering relationships are all areas in which computers currently excel. When you see a computer beat a human on a game show, this is the only form of intelligence that you're actually seeing, out of seven kinds of intelligence. Yes, you might see small bits of other kinds of intelligence, but this is the focus. Basing an assessment of human-versus-computer intelligence on just one area isn't a good idea.

processing used to achieve that goal, and data acquisition used to better understand the goal. AI relies on algorithms to achieve a result that may or may not have anything to do with human goals or methods of achieving those goals. With this in mind, you can categorize AI in four ways:

- **Acting humanly:** When a computer acts like a human, it best reflects the Turing Test, in which the computer succeeds when differentiation between the computer and a human isn't possible (see "The Turing Test" at the Alan Turing Internet Scrapbook for details). This category also reflects what the media would have you believe AI is all about. You see it employed for technologies such as natural language processing, knowledge representation, automated reasoning, and machine learning (all four of which must be present to pass the test). To pass the Turing test, an AI should have all four previous technologies and possibly integrate other solutions (such as expert systems). Mitsuku (found at `https://chat.kuki.ai/` and `http://www.square-bear.co.uk/mitsuku/home.htm`), a chatbot that won the Loebner Prize five times for the most human-like artificial intelligence, is an example of such integration.

TECHNICAL STUFF

The original Turing Test didn't include any physical contact. Harnad's Total Turing Test does include physical contact, in the form of perceptual ability interrogation, which means that the computer must also employ both computer vision and robotics to succeed. Here's a quick overview of other Turing Test alternatives:

- **Reverse Turing Test:** A human tries to convince a computer that that the human is not a computer (for example, the Completely Automatic Public Turing Test to Tell Computers and Humans Apart, or CAPTCHA).
- **Minimum Intelligent Signal Test:** Only true/false and yes/no questions are given.
- **Marcus Test:** A computer program simulates watching a television show, and the program is tested with meaningful questions about the show's content.
- **Lovelace Test 2.0:** A test detects AI through examining its ability to create art.
- **Winograd Schema Challenge:** This test asks multiple-choice questions in a specific format.

Modern techniques include the idea of achieving the goal rather than mimicking humans completely. For example, the Wright Brothers didn't succeed in creating an airplane by precisely copying the flight of birds; rather, the birds provided ideas that led to aerodynamics, which eventually led to human flight. The goal is to fly. Both birds and humans achieve this goal, but they use different approaches.

- **Thinking humanly:** When a computer thinks like a human, it performs tasks that require intelligence (as contrasted with rote procedures) from a human to succeed, such as driving a car. To determine whether a program thinks like a human, you must have some method of determining how humans think, which the cognitive modeling approach defines. This model relies on three techniques:

  - **Introspection:** Detecting and documenting the techniques used to achieve goals by monitoring one's own thought processes.
  - **Psychological testing:** Observing a person's behavior and adding it to a database of similar behaviors from other persons given a similar set of circumstances, goals, resources, and environmental conditions (among other things).
  - **Brain imaging:** Monitoring brain activity directly through various mechanical means, such as Computerized Axial Tomography (CAT), Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), and Magnetoencephalography (MEG).

  After creating a model, you can write a program that simulates the model. Given the amount of variability among human thought processes and the difficulty of accurately representing these thought processes as part of a program, the results are experimental at best. This category of thinking humanly is often used in psychology and other fields in which modeling the human thought process to create realistic simulations is essential.

- **Thinking rationally:** Studying how humans think using some standard enables the creation of guidelines that describe typical human behaviors. A person is considered rational when following these behaviors within certain levels of deviation. A computer that thinks rationally relies on the recorded behaviors to create a guide as to how to interact with an environment based on the data at hand. The goal of this approach is to solve problems logically, when possible. In many cases, this approach would enable the creation of a baseline technique for solving a problem, which would then be modified to actually solve the problem. In other words, the solving of a problem in principle is often different from solving it in practice, but you still need a starting point.

- **Acting rationally:** Studying how humans act in given situations under specific constraints enables you to determine which techniques are both efficient and effective. A computer that acts rationally relies on the recorded actions to interact with an environment based on conditions, environmental factors, and existing data. As with rational thought, rational acts depend on a solution in principle, which may not prove useful in practice. However, rational acts do provide a baseline upon which a computer can begin negotiating the successful completion of a goal.

HUMAN VERSUS RATIONAL PROCESSES

Human processes differ from rational processes in their outcome. A process is *rational* if it always does the right thing based on the current information, given an ideal performance measure. In short, rational processes go by the book and assume that the book is actually correct. Human processes involve instinct, intuition, and other variables that don't necessarily reflect the book and may not even consider the existing data. As an example, the rational way to drive a car is to always follow the laws. However, traffic isn't rational. If you follow the laws precisely, you end up stuck somewhere because other drivers aren't following the laws precisely. To be successful, a self-driving car must therefore act humanly, rather than rationally.

The categories used to define AI offer a way to consider various uses for or ways to apply AI. Some of the systems used to classify AI by type are arbitrary and not distinct. For example, some groups view AI as either strong (generalized intelligence that can adapt to a variety of situations) or weak (specific intelligence designed to perform a particular task well). The problem with strong AI is that it doesn't perform any task well, while weak AI is too specific to perform tasks independently. Even so, just two type classifications won't do the job even in a general sense. The four classification types promoted by Arend Hintze (see "Understanding the four types of AI, from reactive robots to self-aware beings" at Conversation.com for details) form a better basis for understanding AI:

- **Reactive machines:** The machines you see beating humans at chess or playing on game shows are examples of reactive machines. A reactive machine has no memory or experience upon which to base a decision. Instead, it relies on pure computational power and smart algorithms to re-create every decision every time. This is an example of a weak AI used for a specific purpose. (The "Considering the Chinese Room argument" section of Chapter 5 explains the meaning of a weak AI.)
- **Limited memory:** An SD car or autonomous robot can't afford the time to make every decision from scratch. These machines rely on a small amount of memory to provide experiential knowledge of various situations. When the machine sees the same situation, it can rely on experience to reduce reaction time and to provide more resources for making new decisions that haven't yet been made. This is an example of the current level of strong AI.
- **Theory of mind:** A machine that can assess both its required goals and the potential goals of other entities in the same environment has a kind of understanding that is feasible to some extent today, but not in any commercial form. However, for SD cars to become truly autonomous, this level of AI

must be fully developed. An SD car would not only need to know that it must go from one point to another, but also intuit the potentially conflicting goals of drivers around it and react accordingly. (Robot soccer, `http://www.cs.cmu.edu/~robosoccer/main/` and `https://www.robocup.org/`, is another example of this kind of understanding, but at a simple level.)

- **Self-awareness:** This is the sort of AI that you see in movies. However, it requires technologies that aren't even remotely possible now because such a machine would have a sense of both self and consciousness. In addition, instead of merely intuiting the goals of others based on environment and other entity reactions, this type of machine would be able to infer the intent of others based on experiential knowledge.

Understanding the History of AI

The previous sections of this chapter help you understand intelligence from the human perspective and see how modern computers are woefully inadequate for simulating such intelligence, much less actually becoming intelligent themselves. However, the desire to create intelligent machines (or, in ancient times, idols) is as old as humans. The desire not to be alone in the universe, to have something with which to communicate without the inconsistencies of other humans, is a strong one. Of course, a single book can't contemplate all of human history, so the following sections provide a brief, pertinent overview of the history of modern AI attempts.

Starting with symbolic logic at Dartmouth

The earliest computers were just that: computing devices. They mimicked the human ability to manipulate symbols in order to perform basic math tasks, such as addition. Logical reasoning later added the capability to perform mathematical reasoning through comparisons (such as determining whether one value is greater than another value). However, humans still needed to define the algorithm used to perform the computation, provide the required data in the right format, and then interpret the result. During the summer of 1956, various scientists attended a workshop held on the Dartmouth College campus to do something more. They predicted that machines that could reason as effectively as humans would require, at most, a generation to come about. They were wrong. Only now have we realized machines that can perform mathematical and logical reasoning as effectively as a human (which means that computers must master at least six more intelligences before reaching anything even close to human intelligence).

The stated problem with the Dartmouth College and other endeavors of the time relates to hardware — the processing capability to perform calculations quickly enough to create a simulation. However, that's not really the whole problem. Yes, hardware does figure in to the picture, but you can't simulate processes that you don't understand. Even so, the reason that AI is somewhat effective today is that the hardware has finally become powerful enough to support the required number of calculations.

WARNING

The biggest problem with these early attempts (and still a considerable problem today) is that we don't understand how humans reason well enough to create any sort of simulation — assuming that a direct simulation is even possible. Consider again the issues surrounding manned flight described earlier in the chapter. The Wright brothers succeeded not by simulating birds but rather by understanding the processes that birds use, thereby creating the field of aerodynamics. Consequently, when someone says that the next big AI innovation is right around the corner and yet no concrete dissertation exists of the processes involved, the innovation is anything but right around the corner.

Continuing with expert systems

Expert systems first appeared in the 1970s and again in the 1980s as an attempt to reduce the computational requirements posed by AI using the knowledge of experts. A number of expert system representations appeared, including rule based (which use if…then statements to base decisions on rules of thumb), frame based (which use databases organized into related hierarchies of generic information called frames), and logic based (which rely on set theory to establish relationships). The advent of expert systems is important because they present the first truly useful and successful implementations of AI.

TIP

You still see expert systems in use today (even though they aren't called that any longer). For example, the spelling and grammar checkers in your application are kinds of expert systems. The grammar checker, especially, is strongly rule based. It pays to look around to see other places where expert systems may still see practical use in everyday applications.

A problem with expert systems is that they can be hard to create and maintain. Early users had to learn specialized programming languages such as List Processing (Lisp) or Prolog. Some vendors saw an opportunity to put expert systems in the hands of less experienced or novice programmers by using products such as VP-Expert (see *The Illustrated VP-Expert* at Amazon.com), which rely on the rule-based approach. However, these products generally provided extremely limited functionality in using smallish knowledge bases.

In the 1990s, the phrase *expert system* began to disappear. The idea that expert systems were a failure did appear, but the reality is that expert systems were simply so successful that they became ingrained in the applications that they were designed to support. Using the example of a word processor, at one time you needed to buy a separate grammar checking application such as RightWriter. However, word processors now have grammar checkers built in because they proved so useful (if not always accurate; see the *Washington Post* article "Hello, Mr. Chips PCS Learn English" for details).

Overcoming the AI winters

The term *AI winter* refers to a period of reduced funding in the development of AI. In general, AI has followed a path on which proponents overstate what is possible, inducing people with no technology knowledge at all, but lots of money, to make investments. A period of criticism then follows when AI fails to meet expectations, and, finally, the reduction in funding occurs. A number of these cycles have occurred over the years — all of them devastating to true progress.

AI is currently in a new hype phase because of *machine learning,* a technology that helps computers learn from data. Having a computer learn from data means not depending on a human programmer to set operations (tasks), but rather deriving them directly from examples that show how the computer should behave. It's like educating a baby by showing it how to behave through example. Machine learning has pitfalls because the computer can learn how to do things incorrectly through careless teaching.

Five tribes of scientists are working on machine learning algorithms, each one from a different point of view (see the "Avoiding AI Hype and Overestimation" section, later in this chapter, for details). At this time, the most successful solution is *deep learning,* which is a technology that strives to imitate the human brain. Deep learning is possible because of the availability of powerful computers, smarter algorithms, large datasets produced by the digitalization of our society, and huge investments from businesses such as Google, Facebook, Amazon, and others that take advantage of this AI renaissance for their own businesses.

People are saying that the AI winter is over because of deep learning, and that's true for now. However, when you look around at the ways in which people are viewing AI, you can easily figure out that another criticism phase will eventually occur unless proponents tone the rhetoric down. AI can do amazing things, but they're a mundane sort of amazing (such as doing the repetitive work for finding a Covid-19 vaccine; see "How AI is being used for COVID-19 vaccine creation and distribution" at TechRepublic.com). The next section describes how AI is being used now.

Considering AI Uses

You find AI used in a great many applications today. The only problem is that the technology works so well that you don't know it even exists. In fact, you might be surprised to find that many home devices already make use of AI. For example, some smart thermostats automatically create schedules for you based on how you manually control the temperature. Likewise, voice input that is used to control some devices learns how you speak so that it can better interact with you. AI definitely appears in your car and most especially in the workplace. In fact, the uses for AI number in the millions — all safely out of sight even when they're quite dramatic in nature. Here are just a few of the ways in which you might see AI used:

- **Fraud detection:** You get a call from your credit card company asking whether you made a particular purchase. The credit card company isn't being nosy; it's simply alerting you to the fact that someone else could be making a purchase using your card. The AI embedded within the credit card company's code detected an unfamiliar spending pattern and alerted someone to it.
- **Resource scheduling:** Many organizations need to schedule the use of resources efficiently. For example, a hospital may have to determine where to put a patient based on the patient's needs, availability of skilled experts, and the amount of time the doctor expects the patient to be in the hospital.
- **Complex analysis:** Humans often need help with complex analysis because there are literally too many factors to consider. For example, the same set of symptoms could indicate more than one problem. A doctor or other expert might need help making a diagnosis in a timely manner to save a patient's life.
- **Automation:** Any form of automation can benefit from the addition of AI to handle unexpected changes or events. A problem with some types of automation today is that an unexpected event, such as an object in the wrong place, can actually cause the automation to stop. Adding AI to the automation can allow the automation to handle unexpected events and continue as if nothing happened.
- **Customer service:** The customer service line you call today may not even have a human behind it. The automation is good enough to follow scripts and use various resources to handle the vast majority of your questions. With good voice inflection (provided by AI as well), you may not even be able to tell that you're talking with a computer.
- **Safety systems:** Many of the safety systems found in machines of various sorts today rely on AI to take over the vehicle in a time of crisis. For example, many automatic braking systems (ABS) rely on AI to stop the car based on all the inputs that a vehicle can provide, such as the direction of a skid.

Computerized ABS is actually relatively old at 40 years from a technology perspective (see "ABS (Anti-Lock Braking System) — A Brief History Of A 40-Year-Old Life-Saver" at DriveSpark.com for details).

- **Machine efficiency:** AI can help control a machine in such a manner as to obtain maximum efficiency. The AI controls the use of resources so that the system doesn't overshoot speed or other goals. Every ounce of power is used precisely as needed to provide the desired services.

Avoiding AI Hype and Overestimation

This chapter mentions AI hype quite a lot. Unfortunately, the chapter doesn't even scratch the surface of all the hype out there. If you watch movies such as *Her* and *Ex Machina*, you might be led to believe that AI is further along than it is. The problem is that AI is actually in its infancy, and any sort of application such as those shown in the movies is the creative output of an overactive imagination. The following sections help you understand how hype and overestimation are skewing the goals you can actually achieve using AI today.

Defining the five tribes and the master algorithm

You may have heard of something called the singularity, which is responsible for the potential claims presented in the media and movies. The *singularity* is essentially a master algorithm that encompasses all five tribes of learning used within machine learning. To achieve what these sources are telling you, the machine must be able to learn as a human would — as specified by the seven kinds of intelligence discussed in the "Discerning intelligence" section, early in the chapter. Here are the five tribes of learning:

- **Symbologists:** The origin of this tribe is in logic and philosophy. This group relies on inverse deduction to solve problems.
- **Connectionists:** This tribe's origin is in neuroscience, and the group relies on backpropagation to solve problems.
- **Evolutionaries:** The evolutionaries tribe originates in evolutionary biology, relying on genetic programming to solve problems.

- **Bayesians:** This tribe's origin is in statistics and relies on probabilistic inference to solve problems.
- **Analogizers:** The origin of this tribe is in psychology. The group relies on kernel machines to solve problems.

The ultimate goal of machine learning is to combine the technologies and strategies embraced by the five tribes to create a single algorithm (the *master algorithm*) that can learn anything. Of course, achieving that goal is a long way off. Even so, scientists such as Pedro Domingos at the University of Washington are currently working toward that goal.

To make things even less clear, the five tribes may not be able to provide enough information to actually solve the problem of human intelligence, so creating master algorithms for all five tribes may still not yield the singularity. At this point, you should be amazed at just how much people don't know about how they think or why they think in a certain manner. Any rumors you hear about AI taking over the world or becoming superior to people are just plain false.

Considering sources of hype

There are many sources of AI hype out there. Quite a bit of the hype comes from the media and is presented by persons who have no idea of what AI is all about, except perhaps from a sci-fi novel they read once. So, it's not just movies or television that cause problems with AI hype; it's all sorts of other media sources as well. You can often find news reports presenting AI as being able to do something that it can't possibly do because the reporter doesn't understand the technology. Oddly enough, many news services now use AI to at least start articles for reporters (see "Did A Robot Write This? How AI Is Impacting Journalism" at Forbes.com for details).

Some products should be tested a lot more before being placed on the market. The "2020 in Review: 10 AI Failures" article at SyncedReview.com discusses ten products hyped by their developer but which fell flat on their faces. Some of these failures are huge and reflect badly on the ability of AI to perform tasks as a whole. However, something to consider with a few of these failures is that people may have interfered with the device using the AI. Obviously, testing procedures need to start considering the possibility of people purposely tampering with the AI as a potential source of errors. Until that happens, the AI will fail to perform as expected because people will continue to fiddle with the software in an attempt to cause it to fail in a humorous manner.

WARNING

Another cause of problems comes from asking the wrong person about AI. Not every scientist, no matter how smart, knows enough about AI to provide a competent opinion about the technology and the direction it will take in the future. Asking a biologist about the future of AI in general is akin to asking your dentist to perform brain surgery — it simply isn't a good idea. Yet, many stories appear with people like these as the information source. To discover the future direction of AI, it's best to ask a computer scientist or data scientist with a strong background in AI research.

Understanding user overestimation

Because of hype (and sometimes laziness or fatigue), users continually overestimate the ability of AI to perform tasks. For example, a Tesla owner was recently found sleeping in his car while the car zoomed along the highway at 90 mph (see "Tesla owner in Canada charged with 'sleeping' while driving over 90 mph"). However, even with the user significantly overestimating the ability of the technology to drive a car, it does apparently work well enough (at least, for this driver) to avoid a complete failure.

However, you need not be speeding down a highway at 90 mph to encounter user overestimation. Robot vacuums can also fail to meet expectations, usually because users believe they can just plug in the device and then never think about vacuuming again. After all, movies portray the devices working precisely in this manner. The article "How to Solve the Most Annoying Robot Vacuum Cleaner Problems" at RobotsInMyHome.com discusses troubleshooting techniques for various robotic vacuums for a good reason — the robots still need human intervention. The point is that most robots need human intervention at some point because they simply lack the knowledge to go it alone.

Connecting AI to the Underlying Computer

To see AI at work, you need to have some sort of computing system, an application that contains the required software, and a knowledge base. The computing system could be anything with a chip inside; in fact, a smartphone does just as well as a desktop computer for some applications. Of course, if you're Amazon and you want to provide advice on a particular person's next buying decision, the smartphone won't do — you need a really big computing system for that application. The size of the computing system is directly proportional to the amount of work you expect the AI to perform.

The application can also vary in size, complexity, and even location. For example, if you're a business and want to analyze client data to determine how best to make a sales pitch, you might rely on a server-based application to perform the task. On the other hand, if you're a customer and want to find products on Amazon to go with your current purchase items, the application doesn't even reside on your computer; you access it through a web-based application located on Amazon's servers.

The knowledge base varies in location and size as well. The more complex the data, the more you can obtain from it, but the more you need to manipulate it as well. You get no free lunch when it comes to knowledge management. The interplay between location and time is also important. A network connection affords you access to a large knowledge base online but costs you in time because of the latency of network connections. However, localized databases, while fast, tend to lack details in many cases.

IN THIS CHAPTER

- » **Understanding how data analysis works**
- » **Using data analysis effectively with machine learning**
- » **Determining what machine learning can achieve**
- » **Discovering the different kinds of machine learning algorithms**

Chapter 9

Performing Data Analysis for AI

Amassing data isn't a modern phenomenon; people have amassed data for centuries. No matter whether the information appears in text or numeric form, people have always appreciated how data describes the surrounding world, and among other things, they use it to move civilization forward. Data has a value in itself. By using its content, humanity can learn, transmit critical information to descendants (no need to reinvent the wheel), and effectively act in the world.

People have recently learned that data contains more than surface information. If data is in an appropriate numerical form, you can apply special techniques devised by mathematicians and statisticians, called data analysis techniques, and extract even more knowledge from it. In addition, starting from simple data analysis, you can extract meaningful information and subject data to more advanced analytics using machine learning algorithms capable of predicting the future, classifying information, and effectively helping to make optimal decisions.

Data analysis and machine learning enable people to push data usage beyond previous limits to develop a smarter AI. This chapter introduces you to data analysis. It shows how to use data as a learning tool and starting point to solve challenging

AI problems, such as by suggesting the right product to a customer, understanding spoken language, translating English into German, automating car driving, and more.

Defining Data Analysis

The current era is called the Information Age not simply because we have become so data rich but also because society has reached a certain maturity in analyzing and extracting information from that data. Companies such as Alphabet (Google), Amazon, Apple, Facebook, and Microsoft, which have built their businesses on data, are ranked among the most valuable companies in the world. Such companies don't simply gather and keep stored data that's provided by their digital processes; they also know how to make it as valuable as oil by employing precise and elaborate data analysis. Google, for instance, records data from the web in general and from its own search engine, among other things, and in order to support its business, it has built a plurality of machine learning models that are continuously updated based on that data.

You may have encountered the "data is the new oil" mantra in the news, in magazines, or at conferences. The statement implies both that data can make a company rich and that it takes skill and hard work to make this happen. Though many have employed the concept and made it incredibly successful, it was Clive Humbly, a British mathematician, who first equated data to oil, given his experience with consumers' data in the retail sector. Humbly is known for being among the founders of Dunnhumby, a UK marketing company, and the mind behind Tesco's fidelity card program. In 2006, Humbly also emphasized that data is not just money that rains from the sky; it requires effort to make it useful. Just as you can't immediately use unrefined oil because it has to be changed into something else by chemical processes that turn it into gas, plastics, or other chemicals, so data must undergo significant transformations to acquire value.

The most basic data transformations are provided by *data analysis,* and you liken them to the basic chemical transformations that oil undergoes in a refinery before becoming valuable fuel or plastic products. Using just data analysis, you can lay down the foundation for more advanced data analysis processes that you can apply to data. Data analysis, depending on the context, refers to a large body of possible data operations, sometimes specific to certain industries or tasks. You can categorize all these transformations in four large and general families that provide an idea of what happens in data analysis:

- **Transforming:** Changes the data's structure. The term *transforming* refers to different processes, though the most common is putting data into ordered rows and columns in a *matrix format* (also called *flat-file transformation*). For instance, you can't effectively process data of goods bought in a supermarket until you've placed each customer in a single row and added products purchased to single columns within that row. You add those products as numeric entries that contain quantities or monetary value. Transforming can also involve specialized numeric transformations such as *scaling* and *shifting,* through which you change the *mean* (the average) and the *dispersion* (the way a numeric series is spread around its mean values) of the data. These processes make the data suitable for an algorithm.
- **Cleansing:** Fixes imperfect data. Depending on the means of acquiring the data, you may find different problems because of missing information, extremes in range, or simply wrong values. For instance, data in a supermarket may present errors when goods are labeled with incorrect prices. Some data is *adversarial,* which means that it has been created to spoil any conclusions. For instance, a product may have fake reviews on the Internet that change its rank. Cleansing helps to remove adversarial examples from data and to make conclusions reliable.
- **Inspecting:** Validates the data. Data analysis is mainly a human job, though software plays a big role. Humans can easily recognize patterns and spot strange data elements. For this reason, data analysis produces many data statistics and provides useful and insightful visualization, such as Health InfoScape by MIT Senseable Cities and General Electric, which helps grasp informative content at a glance. For example, you can see how diseases connect to one another based on data processed from 72 million records.
- **Modeling:** Grasps the relationship between the elements present in data. To perform this task, you need tools taken from statistics, such as correlations, chi-square tests, linear regression, and many others that can reveal whether some values truly are different from others or just related. For instance, when analyzing expenditures in a supermarket, you can determine that people buying diapers also tend to buy beer. Statistical analysis finds these two products associated many times in the same baskets. (This study is quite a legend in data analytics; see the short story in this *Forbes* article, "Birth of a legend.")

Data analysis isn't magic. You perform transformations, cleansing, inspecting, and modeling by using mass summation and multiplication based on matrix calculus (which is nothing more than the long sequences of summation and multiplication that many people learn in school). The data analysis arsenal also provides basic statistical tools, such as mean and variance, that describe data distribution, or sophisticated tools, such as correlation and linear regression analysis, that reveal whether you can relate events or phenomena to one another (like buying diapers and beer) based on the evidence. To discover more about such data techniques, both *Machine Learning For Dummies,* 2nd Edition, and *Python for Data Science For Dummies,* 2nd Edition, by John Paul Mueller and Luca Massaron (Wiley), offer a practical overview and explanation of each of them.

REMEMBER

What makes data analysis hard in the age of big data is the large volume of data that requires special computing tools, such as Hadoop (`http://hadoop.apache.org/`) and Apache Spark (`https://spark.apache.org/`), which are two software tools used to perform massive data operations. In spite of such advanced tools, it's still a matter of perspiration to manually prepare up to 80 percent of the data. The interesting *New York Times* interview in "For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights" with Monica Rogati, who is an expert in the field and an AI advisor to many companies, discusses this issue in more detail.

Understanding why analysis is important

Data analysis is essential to AI. In fact, no modern AI is possible without visualizing, cleansing, transforming, and modeling data before advanced algorithms enter the process and turn it into information of even higher value than before.

In the beginning, when AI consisted of purely algorithmic solutions and expert systems, scientists and experts carefully prepared the data to feed them. Therefore, for instance, if someone wanted an algorithm to process information, a data expert placed the correct data into *lists* (ordered sequences of data elements) or in other data structures that could appropriately contain the information and allow its desired manipulation. At such a time, data experts gathered and organized the data so that its content and form were exactly as expected, because it was created or selected for that specific purpose. Manipulating known data into a specific form posed a serious limitation because crafting data required a lot of time and energy; consequently, algorithms received less information than is available today.

Today, the attention has shifted from data production to data preparation by using data analysis. The idea is that various sources already produce data in such large quantities that you can find what you need without having to create special data for the task. For instance, imagine wanting an AI to control your pet door to let cats and dogs in but keep other animals out. Modern AI algorithms learn from task-specific data, which means processing a large number of images showing examples of dogs, cats, and other animals. Most likely, such a huge set of images will arrive from the Internet, maybe from social sites or image searches. Previously, accomplishing a similar task meant that algorithms would use just a few specific inputs about shapes, sizes, and distinctive characteristics of the animals, for example. The paucity of data meant that they could accomplish only a few limited tasks. In fact, no examples exist of an AI that can power a pet door using classic algorithms or expert systems.

Data analysis comes to the rescue of modern algorithms by providing information about the images retrieved from the Internet. Using data analysis enables AI to discover the image sizes, variety, number of colors, words used in the image titles, and so on. This is part of inspecting the data and, in this case, that's necessary to cleanse and transform it. For instance, data analysis can help you spot a photo of an animal erroneously labeled a cat (you don't want to confuse your AI) and help you transform the images to use the same color format (for example, shades of gray) and the same size.

Reconsidering the value of data

With the explosion of data availability on digital devices (as discussed in Chapter 2), data assumes new nuances of value and usefulness beyond its initial scope of instructing (teaching) and transmitting knowledge (transferring data). The abundance of data, when provided to data analysis, acquires new functions that distinguish it from the informative ones:

- Data describes the world better by presenting a wide variety of facts, and in more detail by providing nuances for each fact. It has become so abundant that it covers every aspect of reality. You can use it to unveil how even apparently unrelated things and facts actually relate to each other.
- Data shows how facts associate with events. You can derive general rules and learn how the world will change or transform, given enough data to dig out the rules you need.

In some respects, data provides us with new super-powers. Chris Anderson, *Wired*'s previous editor-in-chief, discusses how large amounts of data can help scientific discoveries outside the scientific method (see "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete" at Wired.com). The author relies on the example of achievements of Google in the advertising and translation business sectors, in which Google achieved prominence not by using specific models or theories but rather by applying algorithms to learn directly from data.

DISCOVERING SMARTER AI DEPENDS ON DATA

More than simply powering AI, data makes AI possible. Some people would say that AI is the output of sophisticated algorithms of elevated mathematical complexity, and that's certainly true. Activities like vision and language understanding require algorithms that aren't easily explained in layman's terms and necessitate millions of computations to work. (Hardware plays a role here, too.)

Yet there's more to AI than algorithms. Dr. Alexander Wissner-Gross, an American research scientist, entrepreneur, and fellow at the Institute for Applied Computation Science at Harvard, provided his insights in an earlier interview at Edge.org ("Datasets Over Algorithms"). The interview reflects on why AI technology took so long to take off, and Wissner-Gross concludes that it might have been a matter of the quality and availability of data rather than algorithmic capabilities.

Wissner-Gross reviews the timing of most breakthrough AI achievements in preceding years, showing how data and algorithms contribute to the success of each breakthrough and highlighting how each of them was fresh at the time the milestone was reached. Wissner-Gross shows how data is relatively new and always updated, whereas algorithms aren't new discoveries, but rather rely on consolidation of older technology.

The conclusions of Wissner-Gross's reflections are that, on average, the algorithm is usually 15 years older than the data. He points out that data is pushing AI's achievements forward and leaves the reader wondering what could happen if feeding the presently available algorithms with better data in terms of quality and quantity were possible.

As in advertising, scientific data (such as from physics, chemistry or biology) can support innovation that allows scientists to approach problems without hypotheses, instead considering the variations found in large amounts of data and using discovery algorithms. In the past, scientists took uncountable observations and a multitude of experiments to gather enough deductions to describe the physics of the universe using the scientific method. This manual process allowed scientists to find many underlying laws of the world.

The ability to innovate using data alone is a major breakthrough in the scientific quest to understand the world. AI achivements such AlphaFold (described in "DeepMind solves 50-year-old 'grand challenge' with protein folding A.I." at CNBC.com) allow scientists to figure out how proteins fold in space and how they function without the need for long experimentation. For many other scientific tasks data analysis pairs observations expressed as inputs and outputs. This technique makes it possible to determine how things work and to define, thanks to machine learning, approximate rules (laws) of our world without having to resort to using manual observations and deductions. Many aspects of the scientific process are now faster and more automatic.

Defining Machine Learning

The pinnacle of data analysis is machine learning. You can successfully apply machine learning only after data analysis provides correctly prepared input. However, only machine learning can associate a series of outputs and inputs, as well as determine the working rules behind the output in an effective way. Data analysis concentrates on understanding and manipulating the data so that it can become more useful and provide insights on the world, whereas machine learning strictly focuses on taking inputs from data and elaborating a working, internal representation of the world that you can use for practical purposes. Machine learning enables people to perform such tasks as predicting the future, classifying things in a meaningful way, and making the best rational decision in a given context.

REMEMBER

The central idea behind machine learning is that you can represent reality by using a mathematical function that the algorithm doesn't know in advance, but which it can guess after seeing some data. You can express reality and all its challenging complexity in terms of unknown mathematical functions that machine learning algorithms find and make actionable. This concept is the core idea for all kinds of machine learning algorithms.

Learning in machine learning is purely mathematical, and it ends by associating certain inputs to certain outputs. It has nothing to do with understanding what the algorithm has learned (data analysis builds understanding to a certain extent), thus the learning process is often described as *training* because the algorithm is trained to match the correct answer (the output) to every question offered (the input). (*Machine Learning For Dummies*, 2nd Edition, by John Paul Mueller and Luca Massaron, describes in detail how this process works.)

In spite of lacking deliberate understanding and being simply a mathematical process, machine learning can prove useful in many tasks. It provides the AI application the power of doing the most rational thing given a certain context when learning occurs by using the right data. The following sections help describe how machine learning works in more detail, what benefits you can hope to obtain, and the limits of using machine learning within an application.

Understanding how machine learning works

Many people are used to the idea that applications start with a function, accept data as input, and then provide a result. For example, a programmer might create a function called `Add()` that accepts two values as input, such as 1 and 2, and provide the result, which is 3. The output of this process is a value. In the past, writing a program meant understanding the function used to manipulate data to create a given result with certain inputs. Machine learning turns this process around. In this case, you know that you have inputs, such as 1 and 2. You also know that the desired result is 3. However, you don't know what function to apply to create the desired result. Training provides a learner algorithm with all sorts of examples of the desired inputs and results expected from those inputs. The learner then uses this input to create a function. In other words, training is the process whereby the learner algorithm maps a flexible function to the data. The output is typically the probability of a certain class or a numeric value.

To give an idea of what happens in the training process, imagine a child learning to distinguish trees from other objects. Before the child can do so in an independent fashion, a teacher presents the child with a certain number of tree images, complete with all the facts that make a tree distinguishable from other objects of the world. Such facts could be features such as the tree's material (wood), its parts (trunk, branches, leaves or needles, roots), and location (planted into the soil). The child produces an idea of what a tree looks like by contrasting the display of tree features with the images of other, different objects, such as pieces of furniture that are made of wood but do not share other characteristics with a tree.

A machine learning classifier works the same. It builds its cognitive capabilities by creating a mathematical formulation that includes all the given features in a way that creates a function that can distinguish one class from another. Pretend that a mathematical formulation, also called *target function,* exists to express the characteristics of a tree. In such a case, a machine learning classifier can look for its representation as a replica or an approximation (a different function that works alike). Being able to express such mathematical formulation is the representation capability of the machine learning algorithm.

From a mathematical perspective, you can express the representation process in machine learning by using the equivalent term *mapping.* Mapping happens when you discover the construction of a function by observing its outputs. A successful mapping in machine learning is similar to a child internalizing the idea of an object. The child understands the abstract rules derived from the facts of the world in an effective way so that when the child sees a tree, for example, the child immediately recognizes it.

Such a representation (abstract rules derived from real-world facts) is possible because the learning algorithm has many internal parameters (consisting of vectors and matrices of values), which equate to the algorithm's memory for ideas that are suitable for its mapping activity that connects features to response classes. The dimensions and type of internal parameters delimit the kind of target functions that an algorithm can learn. An optimization engine in the algorithm changes parameters from their initial values during learning to represent the target's hidden function.

During optimization, the algorithm searches the possible variants of its parameter combinations to find one that allows correct mapping between features and classes during training. This process evaluates many potential candidate target functions from among those that the learning algorithm can guess. The set of all the potential functions that the learning algorithm can discover is the *hypothesis space.* You can call the resulting classifier with its set parameters a hypothesis, a way in machine learning to say that the algorithm has set parameters to replicate the target function and is now ready to define correct classifications (a fact demonstrated later).

The hypothesis space must contain all the parameter variants of all the machine learning algorithms that you want to try to map to an unknown function when solving a classification problem. Different algorithms can have different hypothesis spaces. What really matters is that the hypothesis space contains the target function (or its approximation, which is a different but similar function, because in the end all you need is something that works).

You can imagine this phase as the time when a child experiments with many different creative ideas by assembling knowledge and experiences (an analogy for the given features) in an effort to create a visualization of a tree. Naturally, the parents are involved in this phase, and they provide relevant environmental inputs. In machine learning, someone has to provide the right learning algorithms, supply some nonlearnable parameters (called hyperparameters), choose a set of examples to learn from, and select the features that accompany the examples. Just as a child can't always learn to distinguish between right and wrong if left alone in the world, so machine learning algorithms need guidance from human beings to learn successfully.

Understanding the benefits of machine learning

You find AI and machine learning used in a great many applications today. The only problem is that the technology works so well that you don't know that it even exists. In fact, you might be surprised to find that many devices in your home already make use of both technologies. Both technologies definitely appear in your car and the workplace. In fact, the uses for both AI and machine learning number in the millions — all safely out of sight even when they're quite dramatic in nature. Chapter 1 lists a few of the ways in which you might see AI used (fraud detection, resource scheduling, and others; see "Considering AI Uses" in that chapter), but that list doesn't even begin to scratch the surface. You can find AI used in many other ways. However, it's also useful to view uses of machine learning outside the normal realm that many consider the domain of AI. Here are a few uses for machine learning that you might not associate with an AI:

- **Access control:** In many cases, access control is a yes-or-no proposition. An employee smartcard grants access to a resource in much the same way that people have used keys for centuries. Some locks do offer the capability to set times and dates that access is allowed, but such coarse-grained control doesn't really answer every need. By using machine learning, you can determine whether an employee should gain access to a resource based on role and need. For example, an employee can gain access to a training room when the training reflects an employee role.
- **Animal protection:** The ocean might seem large enough to allow animals and ships to cohabitate without problem. Unfortunately, many animals get hit by ships each year. A machine learning algorithm could allow ships to avoid animals by learning the sounds and characteristics of both the animal and the ship.
- **Predicting wait times:** Most people don't like waiting when they have no idea how long the wait will be. Machine learning allows an application to determine waiting times based on staffing levels, staffing load, complexity of the problems the staff is trying to solve, availability of resources, and so on.

Being useful; being mundane

Even though the movies suggest that AI is sure to make a huge splash, and you do occasionally see incredible uses for AI in real life, most uses for AI are mundane and even boring. For example, Hilary Mason, general manager of machine learning at Cloudera, cites how machine learning is used in an international accounting firm to automatically fill in accounting questionnaires (see "Make AI Boring: The Road from Experimental to Practical" at InformationWeek.com). The act of performing this analysis is dull when compared to other sorts of AI activities, but the benefits are that the accounting firm saves money, and the results are better as well.

Specifying the limits of machine learning

Machine learning relies on algorithms to analyze huge datasets. Currently, machine learning can't provide the sort of AI that the movies present. Even the best algorithms can't think, feel, display any form of self-awareness, or exercise free will.

What machine learning can do is perform predictive analytics far faster than any human can. As a result, machine learning can help humans work more efficiently. The current state of AI, then, is one of performing analysis, but humans must still consider the implications of that analysis and make the required moral and ethical decisions. Essentially, machine learning provides just the learning part of AI, and that part is nowhere near ready to create an AI of the sort you see in films.

The main point of confusion between learning and intelligence is people's assumption that simply because a machine gets better at its job (learning), it's also aware (intelligence). Nothing supports this view of machine learning. The same phenomenon occurs when people assume that a computer is purposely causing problems for them. The computer can't assign emotions and therefore acts only upon the input provided and the instruction contained within an application to process that input. A true AI will eventually occur when computers can finally emulate the clever combination used by nature:

- **Genetics:** Slow learning from one generation to the next
- **Teaching:** Fast learning from organized sources
- **Exploration:** Spontaneous learning through media and interactions with others

Apart from the fact that machine learning consists of mathematical functions optimized for a certain purpose, other weaknesses expose the limits of machine learning. You need to consider three important limits:

- **Representation:** Representing some problems using mathematical functions isn't easy, especially with complex problems like mimicking a human brain. At the moment, machine learning can solve single, specific problems that answer simple questions, such as "What is this?" and "How much is it?" and "What comes next?"
- **Overfitting:** Machine learning algorithms can seem to learn what you care about, but they actually often don't. Therefore, their internal functions mostly memorize the data without learning from the data. *Overfitting* occurs when your algorithm learns too much from your data, up to the point of creating functions and rules that don't exist in reality.
- **Lack of effective generalization because of limited data:** The algorithm learns what you teach it. If you provide the algorithm with bad or weird data, it behaves in an unexpected way.

As for representation, a simple-learner algorithm can learn many different things, but not every algorithm is suited for certain tasks. Some algorithms are general enough that they can play chess, recognize faces on Facebook, and diagnose cancer in patients. An algorithm reduces the data inputs and the expected results of those inputs to a function in every case, but the function is specific to the kind of task you want the algorithm to perform.

The secret to machine learning is generalization. However, with generalization come the problems of overfitting and *biased data* (data that when viewed using various statistical measures is skewed in one direction or the other). The goal is to generalize the output function so that it works on data beyond the training examples. For example, consider a spam filter. Say that your dictionary contains 100,000 words (a small dictionary). A limited training dataset of 4,000 or 5,000 word combinations (as you would see them in a real sentence) must create a generalized function that can then find spam in the 2^100,000 combinations that the function will see when working with actual data. In such conditions, the algorithm will seem to learn the rules of the language, but in reality it won't do well. The algorithm may respond correctly to situations similar to those used to train it, but it will be clueless in completely new situations. Or, it can show biases in unexpected ways because of the kind of data used to train it.

For instance, Microsoft trained its AI, Tay, to chat with human beings on Twitter and learn from their answers. Unfortunately, the interactions went haywire because users exposed Tay to hate speech, raising concerns about the goodness of any AI powered by machine learning technology. (You can read some of the story at `https://tinyurl.com/4bfakpac`.) The problem was that the machine learning algorithm was fed bad, unfiltered data (Microsoft didn't use appropriate data analysis to clean and balance the input appropriately), which overfitted the result. The overfitting selected the wrong set of functions to represent the world in a general way as needed to avoid providing nonconforming output, such as hate speech. Of course, even if the output wasn't undesirable, it could still be nonconforming, such as giving wrong answers to straightforward questions. Other AI trained to chat with humans, such as the award-winning Kuki (`https://www.kuki.ai/`), aren't exposed to the same risks as Tay because their learning is strictly controlled and supervised by data analysis and human evaluation.

Considering How to Learn from Data

Everything in machine learning revolves around algorithms. An algorithm is a procedure or formula used to solve a problem. The problem domain affects the kind of algorithm needed, but the basic premise is always the same: to solve some sort of problem, such as driving a car or playing dominoes. In the first case, the problems are complex and many, but the ultimate problem is one of getting a passenger from one place to another without crashing the car. Likewise, the goal of playing dominoes is to win.

Learning comes in many different flavors, depending on the algorithm and its objectives. You can divide machine learning algorithms into three main groups, based on their purpose:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

The following sections discuss what different kinds of algorithms are exploited by machine learning in more detail.

Supervised learning

Supervised learning occurs when an algorithm learns from example data and associated target responses that can consist of numeric values or string labels, such as classes or tags, in order to later predict the correct response when given new examples. The supervised approach is similar to human learning under the supervision of a teacher. The teacher provides good examples for the student to memorize, and the student then derives general rules from these specific examples.

You need to distinguish between *regression problems,* whose target is a numeric value, and *classification problems,* whose target is a qualitative variable, such as a class or a tag. A regression task could determine the average prices of houses in the Boston area, while an example of a classification task is distinguishing between kinds of iris flowers based on their sepal and petal measures. Here are some examples of supervised learning with important applications in AI described by their data input, their data output, and the real-world application they can solve:

Data Input (X)	Data Output (y)	Real-World Application
History of customers' purchases	A list of products that customers have never bought	Recommender system
Images	A list of boxes labeled with an object name	Image detection and recognition
English text in the form of questions	English text in the form of answers	Chatbot, a software application that can converse
English text	German text	Machine language translation
Audio	Text transcript	Speech recognition
Image, sensor data	Steering, braking, or accelerating	Behavioral planning for autonomous driving

Unsupervised learning

Unsupervised learning occurs when an algorithm learns from plain examples without any associated response, leaving the algorithm to determine the data patterns on its own. This type of algorithm tends to restructure the data into something else, such as new features that may represent a class or a new series of uncorrelated values. The resulting data are quite useful in providing humans with insights into the meaning of the original data and new useful inputs to supervised machine learning algorithms.

Unsupervised learning resembles methods used by humans to determine that certain objects or events are from the same class, such as observing the degree of similarity between objects. Some recommender systems that you find on the web in the form of marketing automation are based on this type of learning. The marketing automation algorithm derives its suggestions from what you've bought in the past. The recommendations are based on an estimation of what group of customers you resemble the most and then inferring your likely preferences based on that group.

Reinforcement learning

Reinforcement learning occurs when you present the algorithm with examples that lack labels, as in unsupervised learning. However, you can accompany an example with positive or negative feedback according to the consequences of the solution that the algorithm proposes.

Reinforcement learning is connected to applications for which the algorithm must make decisions (so the product is prescriptive, not just descriptive, as in unsupervised learning), and the decisions bear consequences. In the human world, it is just like learning by trial and error. Errors help you learn because they have a penalty added (cost, loss of time, regret, pain, and so on), teaching you that a certain course of action is less likely to succeed than others. An interesting example of reinforcement learning occurs when computers learn to play video games by themselves.

In this case, an application presents the algorithm with examples of specific situations, such as having the gamer stuck in a maze while avoiding an enemy. The application lets the algorithm know the outcome of actions it takes, and learning occurs while trying to avoid what it discovers to be dangerous and to pursue survival. You can see how Google DeepMind created a reinforcement learning program that plays old Atari video games on YouTube ("Google DeepMind's Deep Q-learning playing Atari Breakout"). When watching the video, notice how the program is initially clumsy and unskilled but steadily improves with training until it becomes a champion. The process is described as having strong and weak points by Raia Hadsell, a senior research scientist on the Deep Learning team at DeepMind, in an enlightening video from TEDx Talks, "Artificial intelligence, video games and the mysteries of the mind," on YouTube.