IN THIS CHAPTER

- » Getting beyond the hype of artificial intelligence (AI)
- » Distinguishing AI from machine learning
- » Understanding the science and engineering in machine learning
- » Delineating where engineering ends and art begins

Chapter 1

Getting the Real Story About AI

Artificial intelligence (AI), the theory and development of computer systems capable of performing tasks that would otherwise require human intelligence, is a vast topic today, and it continues to grow larger all the time, thanks to the constant introduction of new technologies. Despite the complexity of these technologies, most people encounter AI through everyday applications, such as interacting with their digital assistants, receiving shopping recommendations, or creating text, images, and videos to post on social networks using generative AI tools. Talking to your smartphone is both fun and helpful for finding out things like the location of the best sushi restaurant in town or discovering how to get to the concert hall. As you interact with your smartphone, it learns more about the way you talk and makes fewer mistakes in understanding your requests. The capability of your smartphone to comprehend and interpret your unique way of speaking is an example of AI, and it is not the only application available. Part of the technology used to make everything happen is *machine learning*, which involves the use of various techniques to enable algorithms to make predictions based on historical data records.

You also likely encounter and make use of machine learning and AI all over the place today without really noticing. For example, when smart devices adapt to your preferences over time or when digital assistants improve their understanding of your commands, these are examples of machine learning in action. Likewise, recommender systems, such as those found on Amazon, help you decide what to buy based on criteria like previous purchases or products that complement your current choice. The use of both AI and machine learning is expected to keep increasing over time.

In this chapter, you are introduced to AI and machine learning and discover what it means from several perspectives, including how it affects you as a consumer and as a scientist or engineer. You also find that neither AI nor generative AI equals machine learning, even though the media often confuses all the terms. Machine learning is a crucial component of AI, focusing on predicting outcomes based on available information. Generative AI (genAI), which has recently gained importance in the news and our daily lives, is also part of the broader field of AI, but it serves purposes different from predictive machine learning because it aims to create new content, such as text, images, or videos, based on the instructions you provide.

Moving Beyond the Hype

As any technology becomes bigger, so does the hype, and AI certainly has a lot of hype surrounding it. For one thing, some people have chosen to engage in fearmongering rather than science by equating AI with killer robots, such as those depicted in the film *The Terminator*. Actually, your first real experience with a robot is more likely to be in the form of a healthcare assistant or possibly as a coworker. The reality is that you interact with AI and machine learning in far more mundane ways than you might realize.

REMEMBER

You may have also heard more about AI than machine learning. AI is currently receiving the lion's share of attention, but in the form of genAI. As a discipline, AI includes both machine learning and genAI. This chapter helps you understand the relationship between machine learning and AI so that you can better understand how this book enables you to move into a technology that used to appear only within the confines of science fiction novels and films.

Machine learning and AI both have strong engineering components. That is, many aspects of these technologies, particularly the performance and behavior of systems and algorithms, can be measured and optimized through established practices and practical evaluation. In addition, both have strong scientific components, through which researchers test concepts and develop new approaches to simulating or approximating certain aspects of intelligence and decision-making. Finally, machine learning also has an artistic component where intuition, creativity, and experience can play a critical role. This is where a talented practitioner can excel, especially when the results from AI and machine learning may seem counterintuitive, and only the experience and creativity of a skilled practitioner can ensure that models or systems perform as expected.

Dreaming of Electric Sheep

Androids (a specialized kind of robot that looks and acts like a human, such as Data in *Star Trek: The Next Generation*) and some types of *humanoid robots* (a kind of robot that has human characteristics but is easily distinguished from a human, such as C-3PO in *Star Wars*) have become the poster children for AI. They present computers in a form that people can *anthropomorphize* (give human characteristics to, even though they aren't human). In fact, it's entirely possible that one day you won't be able to distinguish between human and artificial life with ease. Science fiction authors, such as Philip K. Dick, have long predicted such an occurrence, and it seems all too possible today. In his novel "Do Androids Dream of Electric Sheep?" Dick discusses the whole concept of more real than real. The idea appears as part of the plot in the movie *Blade Runner*. However, some uses of robots today are just plain fun, as seen with robots serving at restaurants. The sections that follow help you understand how close technology currently gets to the ideals presented by science fiction authors and the movies.

TECHNICAL STUFF

For physical androids, the current state of the art is impressive but still not even close to humans. In text-based interactions, some advanced AI can hold remarkable human-like conversations, but don't be fooled by their linguistic skills, as they lack genuine consciousness or understanding.

Understanding the history of AI and machine learning

There is a reason, other than anthropomorphism, that humans envision the ultimate AI as one that is embodied within some android. Ever since the ancient Greeks, humans have discussed the possibility of placing a mind inside a mechanical body. One such myth is that of a mechanical man called Talos. The fact that the ancient Greeks had complex mechanical devices, of which only one still exists (read about the Antikythera mechanism at `www.ancient-wisdom.com/antikythera.htm`), suggests that their dreams may have been inspired by more than just fantasy. Throughout the centuries, people have discussed mechanical persons capable of thought (such as Rabbi Judah Loew's Golem).

AI is built on the hypothesis that mechanizing thought is possible. During the first millennium, Greek, Indian, and Chinese philosophers all explored formal reasoning and logic, which are the building blocks of the idea of mechanizing thought. As early as the 17th century, Gottfried Leibniz, Thomas Hobbes, and René Descartes discussed the potential for rationalizing all thought as simply mathematical symbols. Of course, the complexity of the problem eluded them. The point is that the vision for AI has been around for an incredibly long time, but the implementation of some working AI is relatively new.

The actual birth of AI as we know it today began with Alan Turing's publication of "Computing Machinery and Intelligence" in 1950 (`https://courses.cs.umbc.edu/471/papers/turing.pdf`). In this paper, Turing explored the idea of how to determine whether machines can think. Of course, this paper led to the Imitation Game involving three players. Player A is a computer, and Player B is a human. Each must convince Player C (a human who can't see either Player A or Player B) that they are human. If Player C can't determine who is human and who isn't in a consistent way, the computer wins.

A persistent issue with AI is excessive optimism. The problem that scientists are trying to solve with AI is incredibly complex. However, the early optimism of the 1950s and 1960s led scientists to believe that the world would produce intelligent machines in as little as 20 years. After all, machines were doing all sorts of amazing things, such as playing complex games. AI currently has its greatest success in areas such as logistics, data mining, advanced natural language processing (conversational AI), advanced computer vision, medical diagnosis, drug discovery, scientific research (for example, protein folding with models like AlphaFold: `https://alphafold.com`), software development, and materials science.

Exploring what machine learning can do for AI

Machine learning relies on algorithms to analyze datasets. Currently, machine learning can't provide the sort of AI that the movies present. Even the best algorithms can't think, feel, present any form of self-awareness, or exercise free will. Machine learning can identify complex patterns, make predictions, and, with generative models, create new data, performing all these tasks far faster than any human can and at a scale that exceeds human capabilities. As a result, machine learning can help humans work more efficiently. A true AI might eventually emerge when computers can finally excel at the clever learning strategies used by nature:

- **Evolution of models and architectures (akin to Genetics):** Slow learning over time, from one generation to the next

- **Supervised Learning (akin to Teaching):** Fast learning from curated sources and explicit guidance
- **Unsupervised Learning (akin to Exploration):** Discovering hidden patterns or intrinsic structures in data
- **Reinforcement Learning (akin to Trial-and-Error):** Learning how to choose the best actions through interaction with an environment and receiving rewards or penalties

The current state of AI, then, is one of performing analysis and suggesting or automating actions, but humans must still consider the implications of that analysis and make the necessary moral and ethical decisions. This is because, as AI systems become more integrated into society and their impact and autonomy increase, it becomes crucial to consider how to use AI responsibly and in a manner that is fair, transparent, accountable, secure, and respectful of human rights. Key considerations include mitigating any bias derived from data or algorithms, ensuring data privacy, enabling human control and oversight, and assessing the societal impacts of any AI system used for automation and decision-making.

REMEMBER

The main point of confusion between learning and intelligence is that people assume 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 instructions contained within an application to process that input.

Considering the goals of machine learning

Currently, AI is based on machine learning, which in turn builds on statistics. Yes, machine learning has a statistical basis, but it makes some different assumptions than statistics do because the goals and approaches are different. Table 1-1 lists some features to consider when comparing machine learning to statistics.

Defining machine learning limits based on hardware

Huge datasets require large amounts of memory. Unfortunately, the requirements don't end there. When you have vast amounts of data and memory, you must also have processors with multiple cores and high speeds. Modern hardware, such as powerful multi-core CPUs or specialized hardware like NVIDIA graphics processing units (GPUs) or Google's tensor processing units (TPUs), enables the massive computational demands of machine learning, especially of deep learning models. Such hardware has architectures that allow parallel computing (performing many calculations simultaneously), which is necessary for handling matrix operations in neural networks. Companies that develop specialized hardware have become central to the AI revolution, such as NVIDIA with its GPUs (including the A100 and H100 series) and software platforms (like CUDA). Other specialized hardware, such as Google's TPUs, which are power-efficient, custom-built chips optimized for speed in running machine learning models, as well as other AI accelerators (AMD, Intel Gaudi, Apple's M-series chips with Neural Engines, and many others), also play a significant role. Without such hardware advancements, the recent breakthroughs in large language models and generative AI would not have been feasible.

TABLE 1-1 Comparing Machine Learning to Statistics

Feature	Machine Learning	Statistics
Data handling	Works with large amounts of structured and unstructured data, aiming at achieving predictive accuracy on unseen data. Consequently, it is crucial to split data into training and test sets.	Methods are focused on hypothesis testing, inference from sample to population, and interpretability.
Data input	The data is sampled, randomized, and transformed to maximize accuracy scoring in the prediction of out-of-sample (or completely new) examples.	Aims to estimate parameters of a population based on an input sample and to quantify the uncertainty of these estimates.
Result	Outputs often include probabilities, scores, or direct predictions, which are used to make informed guesses or decisions.	The output typically includes estimates of parameters, along with measures of uncertainty such as confidence intervals and p-values.
Assumptions	The scientist learns from the patterns in data (data-driven model discovery).	The scientist often starts with a model based on some domain theory.
Distribution	Fewer assumptions are made about the data distribution, or it's learned directly from data.	The scientist tends to assume a well-defined distribution.
Fitting	The scientist creates a best fit, but generalizable, model, having prediction in mind.	The model is fit to the entire sample data to make inferences about the population or to explain relationships within the data.

TECHNICAL STUFF

You will hear more and more often about AI ASICs (Application-Specific Integrated Circuits for Artificial Intelligence). These are specially designed microchips with the type of operations in mind that you need when developing AI technologies, such as matrix multiplications and neural network computations. This specialized design allows them to execute these tasks with maximum efficiency, speed, and lower power consumption. Contrary to CPUs and even GPUs (which can be used for a range of activities), these microchips are not reprogrammable after manufacturing, but they outpace any other solution in the tasks they are designed for.

Apart from the necessary hardware, this book considers some of the following issues as part of making your machine learning experience better:

- **Obtaining a useful result:** As you work through the book, you discover that you need to obtain a valid result first, before you can refine it. In addition, sometimes tuning an algorithm goes too far, and the result becomes quite fragile (and possibly useless outside a specific dataset).
- **Asking the right question:** Many people get frustrated when trying to obtain an answer from machine learning because they keep tuning their algorithm without asking a different question. To use hardware efficiently, sometimes you must step back and review the question you're asking. The question might be wrong, which means that even the best hardware will never find the answer.
- **Relying on intuition too heavily:** All machine learning questions begin as a hypothesis. A scientist uses intuition to create a starting point for discovering the answer to a question. Failure is more common than success when working through a machine learning experience. Your intuition adds the art to the machine learning experience, but sometimes, intuition is wrong, and you have to revisit your assumptions.

Overcoming AI Fantasies

As with many other technologies, AI and machine learning have both their practical and fantasy or fad uses. Of course, the problems with such uses are many. Even if image creation by generative AI has reached an extreme accuracy in generated details and scenery, for one thing, most people wouldn't really want a Picasso or another piece of art created in this manner, except as a fad item (because no one had done it before, or it is trendy to have one for a short time). The point of art isn't in creating an interesting interpretation of a particular real-world representation, but rather in seeing how the artist interpreted it. It was previously believed that computers could only replicate existing artistic styles. However, modern generative AI is increasingly capable of blending styles and producing outputs that can be perceived as novel, though the debate about the true artistic originality of such outputs remains. The following sections discuss AI and machine learning fantasies of various sorts.

Discovering the fad uses of AI and machine learning

AI is entering an era of innovation that you once only read about in science fiction. Given the hype surrounding the idea and possible applications of AI, it can be hard to determine whether a particular AI use is real or simply the dream child of a scientist. The fact is that AI and machine learning will both present potential opportunities to create something amazing, and that we're already at the stage of creating some of those technologies, but you still need to take what you hear with a huge grain of salt.

REMEMBER

To make the future uses of AI and machine learning match the concepts that science fiction has presented over the years, real-world programmers, data scientists, and other stakeholders need to create tools. Nothing happens by magic, even though it may look like magic when you don't know what's happening behind the scenes. In order for the fad uses for AI and machine learning to become real-world uses, developers, data scientists, and others need to continue building real-world tools that may be hard to imagine at this point.

Considering the true uses of AI and 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 most especially in 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 pretty 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 need to determine where to place 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 unforeseen event, such as an object in the wrong place, can actually cause the automation to stop. Adding AI to the automation can enable it to handle unexpected events and continue as if nothing had 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 rely on AI to stop the car based on all the inputs that a vehicle can provide, such as the direction of a skid.
- **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.

This list doesn't even begin to scratch the surface. You can find AI used in many other ways. However, beyond the widely publicized AI applications, machine learning powers many specific, often less visible, solutions that might not always be immediately associated with the usual concept of AI in popular discourse. Here are a few uses for machine learning that you might not commonly 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 when access is allowed, but the 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's role.

- **Animal protection:** The ocean might seem large enough to allow animals and ships to coexist without problem. Unfortunately, many animals get hit by boats 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 of 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 and being mundane

Even though the movies make it sound like AI is going to make a huge splash, and you do sometimes see some incredible uses for AI in real life, the fact of the matter is that most uses for AI are mundane, even boring. The act of performing this analysis using AI is dull when compared to other sorts of AI activities, but the benefits, such as cost savings and improved results from AI-driven analysis, are substantial.

First, Python developers (see Chapter 6 for an overview of the Python language) have developed a vast array of libraries available to make machine learning easy and effective. The Python open-source community is particularly active in creating libraries that make the development of complex machine learning applications accessible to everyone, as seen with the machine learning library Scikit-learn (`https://scikit-learn.org/stable`). In addition, numerous resources are available, such as Kaggle (`www.kaggle.com`), which offers competitions that enable machine learning developers and practitioners to refine their machine learning skills in creating practical applications. The results of these competitions often appear later as part of products that people actually use.

Considering the Relationship Between AI and Machine Learning

Machine learning is only part of what a system requires to become an AI. The machine learning portion of the picture enables an AI to perform these tasks:

- Adapt to new circumstances that the original developer didn't envision
- Detect patterns in all sorts of data sources
- Create new behaviors based on the recognized patterns
- Make decisions based on the success or failure of these behaviors

The use of algorithms to manipulate data is the centerpiece of machine learning. To be successful, a machine learning application must use an appropriate algorithm to achieve a desired result. In addition, the data must lend itself to analysis using the desired algorithm, or it requires careful preparation by scientists.

AI encompasses many other disciplines to simulate the thought process successfully. In addition to machine learning, AI usually includes

- **Natural language processing:** The act of allowing language input and putting it into a form that a computer can use.
- **Natural language understanding:** The act of deciphering the language in order to act upon the meaning it provides.
- **Natural language generation:** The act of creating meaningful language outputs to communicate with humans.
- **Knowledge representation:** The ability to store information in a form that makes fast access possible.
- **Planning (in the form of goal seeking):** The ability to use stored information to draw conclusions in *near real time* (almost at the moment it happens, but with a slight delay, sometimes so short that a human won't notice, but the computer can).
- **Perception and action (often employed in Robotics):** The ability to perceive the environment and act upon it, sometimes in a physical form.

In fact, you might be surprised to find that the number of disciplines required to create an AI is huge. Consequently, this book exposes you to only a portion of what an AI contains. However, even the machine learning portion of the picture can become complex because understanding the world through the data inputs that a computer receives is a complex task. Just think about all the decisions that you constantly make without thinking about them. For example, just the concept of seeing something and knowing whether you can interact successfully with it can become a complex task.

Considering AI and Machine Learning Specifications

As scientists continue to work with technology and turn hypotheses into theories, the technology becomes more related to *engineering* (where theories are implemented) than *science* (where theories are created). As the rules governing a technology become clearer, groups of experts work together to define these rules in written form. The result is a set of *specifications* (a group of rules that everyone agrees upon).

Eventually, implementations of the specifications become *standards* that a governing body, such as the IEEE (Institute of Electrical and Electronics Engineers) or a combination of the ISO/IEC (International Organization for Standardization/International Electrotechnical Commission), manages. Although the field is still rapidly developing, AI and machine learning have both been around long enough to have established standards, such as methodologies, benchmarks, and frameworks, for development and risk management. Numerous domain-specific standards and influential frameworks have emerged, such as those from organizations like the National Institute of Standards and Technology (NIST) with its AI risk framework (`www.nist.gov/itl/ai-risk-management-framework`).

REMEMBER

The basis for machine learning is math. Algorithms determine how to interpret data in specific ways. The mathematical basics for machine learning are presented in Part 3 of this book. You discover that algorithms process input data in specific ways and create outputs by learning from data patterns, which are then used to make predictions or generate new content. What isn't predictable is the data itself. The reason you need AI and machine learning is to decipher the data in a way that allows you to identify patterns and make sense of them.

You can see the details of various algorithms in Part 4, which outlines the algorithms used to perform specific tasks. When you get to Part 5, you begin to see the best practices, common approaches, and established methods when using algorithms to perform tasks. The point is to use an algorithm that will best suit the data you have at hand to achieve the specific goals you've created. Professionals implement algorithms using programming languages that are best suited for the task. Machine learning relies on Python and R, as well as, to some extent, MATLAB, Java, Julia, and C++.

Defining the Divide Between Art, Science, and Engineering

AI and machine learning are considered to be, at the same time, scientific disciplines, engineering fields, and even art forms for good reasons. First, there are the scientific aspects that guide the research focused on machine learning. The scientific elements of AI involve hypothesis testing, experimentation, and the discovery of knowledge to be applied in solving practical problems. This brings us to the engineering aspect, as building machine learning systems that work effectively and solve problems requires software engineering, system design, and optimization of these systems.

The artistic element of machine learning, instead, takes many forms. Choosing the proper data, features, models, and hyperparameters often requires intuition and experience. There are no specific step-by-step instructions, like cooking recipes, to follow to obtain your result. Every problem presents different challenges and multiple acceptable solutions. You can only experiment creatively and iterate numerous times, looking for your way to solve the problem using machine learning algorithms, guided by intuition and sometimes ingenuity.

REMEMBER

Even trivial activities related to machine learning, such as data cleaning, can involve an element of judgment and experience that influences the outcome. How a scientist prepares the data for use is the key. Some tasks, such as removing duplicate records, occur regularly. However, a scientist may also choose to filter the data in some ways or look at only a subset of the data. As a result, the cleaned dataset used by one scientist for machine learning tasks may not precisely match the cleaned dataset used by another.

As a practitioner, you can also tune the algorithms in specific ways or, in this case, more as a researcher, refine how the algorithm works. Again, the goal is to generate output that reveals the desired patterns, allowing you to make sense of the data. For example, when analyzing a picture, a machine learning algorithm must determine which elements of the image reveal its contents and which elements are irrelevant. The answer to that question is crucial if the algorithm is to correctly classify the elements in the image and achieve specific goals.

When working in a machine learning environment, you also have the problem of input data to consider. For example, the microphone found in one smartphone won't produce precisely the same input data that a microphone in another smartphone will. The characteristics of the microphones differ, yet the result of interpreting the vocal commands provided by the user must remain the same. Likewise, environmental noise changes the input quality of the vocal command, and the smartphone can experience certain forms of electromagnetic interference. Clearly, the variables that a designer faces when creating a machine learning environment are both large and complex.

The art behind the engineering is an essential part of machine learning. The experience that a scientist gains in working through data problems is essential because it provides the means for the scientist to make informed choices that make the algorithm work better. A finely tuned algorithm can make the difference between a robot successfully threading a path through obstacles and hitting every one of them.

Predicting the Next AI Winter

The development of machine learning and AI has become increasingly relentless and unstoppable for several reasons, including the availability of more powerful hardware and vast amounts of data (much of it generated by the Internet) to feed algorithms. Businesses, however, don't care about the progress alone; they are looking for new ways to generate cash quickly based on these technologies. Obviously, that's not easy because new technologies like AI do not fit neatly into the existing framework since they are disruptive and so new that you lack guidance and experience on how to gain profit from them. Developer-entrepreneurs exacerbate the problem by overselling technologies. They suggest that the state of the art is more advanced than it actually is, often to secure funding, increase their influence, or advance their careers. In the past, because of the difference between timing and expectations, machine learning and AI have both experienced *AI winters*, a period of time when businesses show little to no interest in developing new processes, technologies, or strategies. At this very moment, the winds are in favor of AI development, as investments and resources are being poured in, and there is considerable excitement in every industry about the potential returns this technology promises. However, a sudden cool-off may always be around the corner, possibly caused by some unmet expectation that could throw investors into disillusionment. Other slowdowns may also contribute, driven by external factors such as climate change, economic downturns, or geopolitical risks.

The first AI winter occurred due to unfulfilled expectations stemming from the overselling of the technology and unexpected difficulties. During the summer of 1956, various scientists attended a workshop held on the campus of Dartmouth College to develop artificially intelligent machines. They predicted that machines that could reason as effectively as humans would require, at most, a generation to come about. They were wrong. Only recently, large language models, such as ChatGPT or Google Gemini, have shown surprising abilities in complex tasks, sometimes mimicking aspects of mathematical and logical reasoning. However, to achieve true human understanding, as Yann LeCun (he is Meta's chief AI scientist, Turing Award winner, NYU data scientist, and one of the pioneers of artificial intelligence) has argued, an AI would also need to demonstrate intelligence in reasoning, as well as in the visual-spatial, bodily-kinesthetic, creative, interpersonal, intrapersonal, and linguistic realms. The stated problem with Dartmouth College and other endeavors of the time was often attributed 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 into the picture, but you can't simulate processes that you don't understand, especially if you lack suitable data and algorithms. Even so, the reason that AI is somewhat effective today is that the hardware has finally become powerful enough to perform the required number of calculations and support improved algorithms and available data.

REMEMBER

While some argue that current AI paradigms might be approaching the limits of their existing architectures and that we shouldn't expect any further breakthroughs, the field is simultaneously experiencing an explosion of new applications and incremental improvements, particularly driven by large-scale models and generative AI. The debate continues as to whether this rapid expansion is sustainable or if inflated expectations in some areas could lead to disillusionment.

Some foresee an AI winter in the near future, partly because the terms *machine learning, deep learning,* and *AI* have become pervasively overused, sometimes in ill-defined ways that can inflate expectations. Undoubtedly, the recent GenAI deluge has accelerated the adoption of AI solutions and demonstrated how AI can power both background processes and customer interactions through sophisticated chatbots and personalized content creation. However, a hype correction may occur if businesses begin to critically scrutinize the return on investment (ROI) of AI solutions or realize that such solutions are only marginally better than simpler automations, albeit at a higher cost and complexity (which is often referred to as technical debt). In such a case, a broader disillusionment could set in, investment will decrease, and a new AI winter may ensue. In many cases, industry observers and scientists, who are concerned with an upcoming AI winter, advocate for a focus on the real, tangible, and demonstrable achievements of AI and machine learning today, rather than continuing to hype some nebulous, far-future capabilities that risk overpromising and underdelivering.

Before you get the idea that everyone is expecting another AI winter, you need to look at the other side of the argument. Some are saying that machine learning and AI have become so deeply embedded that an AI winter really isn't possible anymore. Typically, industry reports often highlight that machine learning and AI haven't met certain goals, such as creating autonomous vehicles or replacing workers in specific tasks. Even though these goals aren't feasible today, the potential exists for achieving them in the future when scientists have completed more research. Moreover, due to the research conducted and the applications created, both machine learning and AI have become profitable in certain fields, so businesses will continue to support them.

When considering the future of machine learning and AI, adopting a more moderate approach is likely the best course of action. At this point, data scientists and other researchers need to take a step back and consider the next level. The current technologies can only take us so far. They're profitable, but, for instance, they can't yet produce a fully autonomous self-driving car that can, without any human intervention, handle difficult weather conditions and unexpected driving scenarios. Also, the current technologies certainly can't produce a robot of the intelligence found in the film *Ex Machina.* There's so much more work to be done, yet. So, if there is an AI winter, it's likely to be a mild one because companies like `Amazon.com` and Google aren't going to throw their technologies out because a few reporters think that they should. In short, the concepts, ideas, and technologies that you discover in this book remain viable and allow you to move forward in a career of your choice.