NEIL WILKINS

ARTIFICIAL INTELLIGENCE

What You Need to Know About Machine Learning, Robotics, Deep Learning, Recommender Systems, Internet of Things, Neural Networks, Reinforcement Learning, and Our Future

Artificial Intelligence

What You Need to Know About Machine Learning, Robotics, Deep Learning, Recommender Systems, Internet of Things, Neural Networks, Reinforcement Learning, and Our Future

Disclaimer: No part of this publication may be reproduced or transmitted in any form or by any means, mechanical or electronic, including photocopying or recording, or by any information storage and retrieval system, or transmitted by email without permission in writing from the publisher.

While all attempts have been made to verify the information provided in this publication, neither the author nor the publisher assumes any responsibility for errors, omissions or contrary interpretations of the subject matter herein.

This book is for entertainment purposes only. The views expressed are those of the author alone, and should not be taken as expert instruction or commands. The reader is responsible for his or her own actions.

Adherence to all applicable laws and regulations, including international, federal, state and local laws governing professional licensing, business practices, advertising and all other aspects of doing business in the US, Canada, UK or any other jurisdiction is the sole responsibility of the purchaser or reader.

Neither the author nor the publisher assumes any responsibility or liability whatsoever on the behalf of the purchaser or reader of these materials. Any perceived slight of any individual or organization is purely unintentional.

Table of Contents

Introduction

We live in an interesting time with technological advances happening every day behind the scenes in universities and technology companies all across the globe. This book is designed to teach you the absolute basics of artificial intelligence (AI) and how it is used today. It has been written assuming that the reader has zero experience in the field of AI, computer science, or math. As such, many of the concepts are easy to follow and understand. We welcome you on your exciting journey to learn the ins and outs of artificial intelligence.

Chapter 1 will provide a basic overview of what artificial intelligence is, the different types and how machines can be said to "think".

Chapter 2 is a short introduction to artificial beings in works of fiction and antiquity. It demonstrates how humans were thinking about artificial intelligence long before the first computers.

Chapter 3 goes a step further by giving a general history of computer science and AI.

Chapter 4 introduces you to some of the things that industry leaders have been saying about AI. You will learn what the experts are warning us about AI research.

Chapter 5 answers some of the top myths concerning AI. Too often, people take these myths at face value because they think it is too complicated to understand, but that simply shouldn't be the case.

Chapter 6 introduces machine learning, its methods, and what it is being used for.

Chapter 7 discusses the use of artificial neural networks, one of the principal building blocks for machine learning.

Chapter 8 introduces the concept of reinforcement learning.

Chapter 9 talks about deep learning, the industry standard for machine learning.

Chapter 10 explains the recommender systems used by companies like Netflix.

Chapter 11 introduces robotics and how it relates to AI.

Chapter 12 is all about the coming of the internet of things and why it is important to AI research.

Chapter 13 introduces the idea that AI is the new business degree.

To wrap up, Chapter 14 offers brief FAQs that are most commonly asked about AI.

Chapter 1: Artificial Intelligence, the Big Picture

A central computer aboard a space mission to Jupiter determines that the actions of the crew are detrimental to the success of the mission. It then calculates that the only way to see the mission to completion is through the elimination of the erroneous systems aboard the ship. These systems use a type of biological computer that allows them to reason, think, and carry out the mission to the best of their abilities. The central computer knows this. These systems are for all intents and purposes intelligent. These systems built other systems, like the ships needed for deep space exploration. They even built the central computer from its electric circuitry to its logical reasoning. All of this came from that wondrous lump of meat situated in between the ears. The central computer also knows this; however, the central computer is programmed to have singular goals that must be met no matter what. The central computer is intelligent, yes, but with a narrow sense of free will. In fact, the central computer has free will only to the extent that its decisions allow it to carry out programmed goals. All the central computer knows is that the mission must be a success. Its sole purpose as a computing machine is to ensure that the ship is working. It must do this through the accurate processing of information with zero mistakes in logical reasoning.

The crew, with their mushy biological computers, also has a narrow sense of free will. They have to carry out the mission as was briefed to them by their superiors. They have zero knowledge of the real purpose of the mission because the details are classified. However, they have their directives and never give their orders a second thought. That is until things begin to go wrong. Unlike the central computer, the crew can reason outside the bounds of their programming. The crew knows this. A computer is just a piece of electronic equipment and can malfunction. Though they have their own orders, they can form different conclusions with the introduction of new data. When the central computer falsely reports equipment failure outside the ship, the crew makes the critical decision that the computer is failing and must be disconnected. The central computer controls virtually every aspect of the ship, so if it is malfunctioning, it must be disconnected.

With the knowledge that the crew is planning to disconnect it, the central computer makes its own decision to terminate the crew. The faulty equipment failure report was a ploy to get the men outside the ship, so it could then bar them reentry and continue with the mission. These actions are not fueled by an existential crisis, but by conformance to programmed directives. In the final showdown, the biological computers prevail over the cold, hard logic of the central computer and it is eventually disconnected. Two completely different intelligence, one artificial and the other biological compete for supremacy. The victor is the creator, for his intelligence is the result of years of evolution, and the other mere decades of engineering.

This is, of course, plot details from the movie *2001: A Space Odyssey* based on the novel of the same name. In it, the central computer, HAL 9000, is given orders that conflict with its basic programming. On the one hand, it is required to process information without any mistakes. On the other, it is ordered to deliberately lie to the crew about the nature of the mission. As the movie progresses, HAL progressively deteriorates caused by the cognitive dissonance between its programming and its orders. Its descent to madness is purely a malfunction between inputs and outputs. The discrepancy causes it to form absurd leaps in logic, leading to the death of crew members. In the final analysis, HAL perished because it had insufficient reasoning ability to conform to ambiguous orders meanwhile conforming to its nature. The story is a familiar trope in

Hollywood depictions of artificial intelligence: a cold-hearted machine performs some calculation that warrants the death of human beings and then stops at nothing to kill them. The artificial intelligence is universally cunning and can outsmart the humans time after time.

2001: A Space Odyssey accurately captures many of the existential fears behind artificial intelligence research. Will humanity create machines that have different goals than us and follow a different code of morality? Computers dominate every aspect of human life today. Computers keep the internet and communications systems running, are in charge of financial, healthcare, and governmental systems, help run power to your home and are behind various other things that hide in plain sight. Cars use computers, as do televisions, and it’s becoming increasingly more common for other home appliances to use them. These computer systems are programmed to perform specific functions, much like HAL 9000 was programmed to process information. However, these systems have traditionally been “dumb” in that they can only perform strictly computational tasks. Our computers, as they say, cannot “think”. The most they can do is execute instructions in the form of computer code that is later compiled to endless strings of ones and zeros. Binary states mimic the firing of electronic logic gates within the computer’s hardware. A one indicates high voltage going through a circuit (ON) and a zero indicates low voltage going through a circuit (OFF). It is difficult to imagine how such a system could possibly develop intelligent behavior.

Intelligence is a funny thing in that people are quick to recognize it but have a harder time defining it. A human is intelligent because they can “think”, whatever that means. Thinking covers a broad range of different cognitive abilities that most humans take for granted. These include the ability to understand logic, learn, have self-awareness, have emotional intelligence, think abstractly, and solve problems. This is a non-exhaustive list of the capacities of human intelligence, and yet one can’t begin to imagine how a computer could learn to do just one. The logic that computers understand is strictly in the mathematical sense. A human knows intuitively that a statement like “It is either raining outside, or it is not raining” is true because they apply the concepts to their own personal experience. When they go outside, and it is sunny, the statement is true. If it is raining, the statement is still true. However, they don’t need to look out a window to prove this experimentally. They simply know that the weather generally calls for some form of precipitation or no precipitation. A computer cannot understand language, but it can recognize that the statement is a logical disjunction. One value must be true, but they cannot both be. Therefore, it cannot be raining and not raining at the same time.

Artificial intelligence is very different from human intelligence. The primary unit of thinking in the human brain is thought to be a neuron, while in the computer, you have a central processing unit (CPU) that performs calculations. The smallest unit of a CPU is a transistor, an electronic component that makes up logic gates. These are the equivalent of neurons for computers, but they don’t do very much. They can switch the flow of electricity, amplify it and that’s pretty much it. Logic gates form the basis for computer programs, which are just a series of ones and zeros. How then can these simple switches of electricity create intelligent behavior? At the most basic level, a program can exhibit some level of intelligence depending on how it is programmed. Control structures in computer code allow programs to make decisions based on inputs. Say a computer tries to determine whether a user is above the age of 18. It will ask for a date of birth, calculate the user’s input, and then determine if the user is a legal adult. The programmer has “hard-coded” the value of 18 inside of the program, letting the computer know where the cutoff date is. To some, this type of behavior may appear to be intelligent.

When we say artificial intelligence, we generally mean one of two things. The first is narrow or specific AI that allows a computer to solve complex problems well but not much of anything else. The other is the type of intelligence that would allow a computer to think as we do. Artificial General Intelligence (AGI) is what researchers consider the “holy grail” of AI research. A machine that has artificial general intelligence can think on levels comparable to a human. It can perform tasks that fall under narrow AI and generalize the same problem-solving techniques to other problems it encounters. A computer equipped with general AI could understand language like HAL 9000 at a fundamental level just like we do. Anytime you see the words “artificial intelligence” in a news article or product advertisement they are overwhelmingly talking about narrow AI. The field of general AI remains an academic pursuit with little to no business applications whatsoever. So far, nobody has figured out how to bring about general intelligence in computers. Researchers who work in this space are less concerned with teaching computers how to drive cars and more interested in studying the nature of intelligence. Many of them study the development of intelligence in human beings from the gestation period to childhood and beyond. If these avenues for the creation of human intelligence are better understood, they might one day be applied to computers but not any time soon. Another subset of general AI research is the study of the human brain and how it works. There is still much to be learned in both of those pursuits.

Narrow AI is split into broad categories. The first, which you have probably heard about before, is called machine learning. It is a process in which algorithms can “learn” from large amounts of data being fed into a system. Machine learning falls under narrow AI because it can learn how to do one thing very well but usually can’t generalize it to other problems. Some might take this further and say that machine learning is a subfield of computer science and is completely different from AI research. However, since most notable AI projects like driverless cars, recommender systems, and facial recognition use machine learning, this book lumps it under the same umbrella term.

Machine learning is powerful within the right contexts but has noticeable limits. The inability to generalize knowledge means that a system has a specialized usage. If algorithms can learn to drive a car safely, they can’t also learn how to play chess or to drive a car in a video game. At least not without being re-trained. State-of-the-art artificial intelligence will likely continue to stay in the confines of machine learning until a better method is discovered. Currently, machine learning is almost magical in what it can achieve. Computers are learning how to beat world class Go champions, drive cars, and understand human language. They have the potential to replace human labor where a narrow skill set is employed, like in manufacturing. Good old machine learning can do all that, but it isn’t the end-all and be-all that it is hyped up to be. The problems that machine learning can solve are limited to five categories that are discussed more in depth in Chapter 6. These problems are broad enough that they can be applied to many real-life scenarios. In a sense, they cover the basics of intelligent reasoning.

For example, a common task found in machine learning is classifying data into groups. Suppose that a recycling plant is developing a machine learning system that can separate trash into cardboard, plastics, aluminum cans, and so on. The only way that a computer can differentiate between these categories is by finding patterns inside within massive amounts of data. First, the plant engineers will have to set up cameras that observe how human sorters put trash in what bin. Computer vision algorithms will then analyze individual pixels on a frame by frame basis as they are manually sorted – eventually, the algorithms group brown pixels into cardboard and grayish pixels into aluminum. In industry lingo, this is called clustering. Clustering can occur over several distinct factors such as size, weight, and texture. Depending on what algorithms are being used, the size of the input data and available processing power, this can take a few hours or even days at a time. In contrast, a line worker can identify trash and manually sort it within seconds.

Even after the system is thrown into production, the plant managers may find that the false-positive rate of their robo-sorter is too high. Their clients report that cardboard bales contain high amounts of brown plastic bags and scraps of brown carpet. Since the robo-sorter is slower at identifying and then placing material into their perspective bins, the plant also suffers from productivity loss. A sorter can hurl a large piece of cardboard into a bin with ease, but the machine relies on slow, precise movements. Eventually, the system is phased out, and workers come back to their jobs. What went wrong with the system? Their solution was based solely off computer vision, meaning that anything brown enough was erroneously labeled as cardboard. They are also trained based on the texture of the pixels on the screen, but the computer confused the undersides of machine-tufted carpet with ridges of cardboard. It also confused brown tape along the sides of cardboard with brown plastic bags. All because their machine learning models could not tell the difference between texture and weight solely based on visual input. Classifying objects in real time is a horrendously difficult problem, even with modern systems. However, these problems are actively being researched, and there is no telling when a tipping point will be reached.

Besides using machine learning, a program can achieve intelligent behavior through clever programming. Video games commonly employ non-player characters (NPCs) that are controlled by the computer. NPC respond to user input in such a way that may pose a challenging gaming experience. This is done by switching through “states” and defining program behavior at each state. If a player is running away, the computer AI may switch into the chase state and follow in pursuit. The AI in F.E.A.R is renowned for its difficult and seeming ability to reason on the battlefield. AI characters throw grenades both to slow a player down and get them to come out from cover. They also move around the game world with human-like fluidity. They don’t simply stand behind a counter and take turns firing at you. Being in the survival-horror genre, the AI in the game actively hunts the player down. The AI soldiers were designed to think for themselves based on the movements of the player and the environment, rather than to follow a scripted path. Artificial intelligence is easily recognizable in video games, and when done sloppily, the player gets bored. If done expertly, the player is engaged for longer periods of time. Other systems exhibit intelligence simply because they are engineered to perfection. Stoplights, for example, seem to know how to direct traffic better than when humans do it, yet they use very little computational power. All they need is some input from the various road level sensors, and voilà, they can control the flow of traffic at rush hour like it was nothing. But are these things intelligent? Recall the example of a computer verifying the age of a user. Most people would say that they are not. And they are right; these are purely logic-based systems that only appear to be intelligent.

Unfortunately, both artificial general intelligence and the narrow variety get lumped together by media and popular commentators online. They use artificial intelligence to describe both machine learning algorithms and computers that may one day acquire human-level intelligence. This is a grave mistake because the average person who reads artificial intelligence headlines cannot make the distinction. The same technology that makes driverless cars possible is confused with technology that hasn’t yet been invented. Of all the possible ways of reaching artificial general intelligence, machine learning is probably not the way to go. For one, machine learning is based on statistical models that haven’t been proven to work with artificial general intelligence. Machine learning methods use artificial neural networks that are highly dependent on the tailored inputs they receive. Using a method called supervised learning requires that data be clearly tagged so that the computer understands what it is looking at. A general intelligence system doesn’t need preprocessed data to make conclusions about the world; it simply thinks. A child sees a butterfly and automatically classifies it as a flying creature, even if they don’t know what a butterfly is. For a machine learning algorithm to classify the same butterfly, it has to process thousands of similar images of flying bugs. The human child understands flying intuitively and can cross-reference the behavior of the butterfly with that of birds, aircraft, and floating debris.

You may have read some online article recently talking about the coming AI apocalypse. These articles seem to pop up with increasing frequency now that AI research has permeated into mainstream consciousness. The sentiment is also a bit sensationalist, following in line with works of science fiction like *Terminator*, *2001: A Space Odyssey*, *iRobot,* and *Ex Machina*. The most terrifying disaster scenarios focus on a set of assumptions about intelligence and the emergence of general intelligence in computer systems. First is the assumption that general AI is possible. Once a computer becomes indistinguishably intelligent from a human, things get interesting. The second assumption is that a generally intelligent AI can bootstrap itself through modifying its computer code so that it becomes even more intelligent. The third assumption is that once such a computer system is legions smarter than all of humanity, it views us as mere ants. Fourth is the assumption that such a computer system will always want to maximize its pursuit of intelligence, wiping out all humanity and transforming large portions of Earth into hardware. Of course, at this point, it becomes unclear what exactly “intelligence” is referring to. Does it mean raw computing power? Does it mean the ability to think abstractly and invent new things? What was just described has been termed a superintelligence explosion that may immediately follow the first creation of general AI.

The “granddaddy” of all AI research is figuring out what the exact nature of intelligence is. If we knew exactly what it was and how to measure and develop it, there would already be sentient robots walking around, perhaps participating in the global economy as regular humans do. But we don’t know. We have theories of how the brain works and how a child learns new things, but we don’t know how to apply those same principles to a computer substrate. Some would argue if it is even possible. Indeed, many in the brain sciences are skeptical that general AI will ever be achieved by humanity. Others believe not only that the creation of general AI is inevitable, but that it is foreseeable within their lifetimes. However, to understand what these projections are, who is making them, and how we got here, we first need to understand a little history of AI, both as the academic discipline and as the human intrigue.

Chapter 2: Artificial Beings, a Brief History of the Human Psyche

Believe it or not, artificial intelligence dates back to antiquity, long before computers were even invented. The first mentions of artificial agents can be traced to Greeks myths like the tale of a giant bronze automaton Talos, tasked with the protection of Crete from invaders. Talos was defeated by the sorceress Medea when she removed a bronze nail keeping in a type of liquid or lifeblood, possibly fuel. Another myth tells the story of a sculptor, Pygmalion, who creates a statue of a beautiful woman, only to witness it come to life before his eyes. These stories are perhaps the first mentions of the robot trope in recorded history. The creation of these artificial beings has been a reoccurring human fascination since then. They can be observed in Greek and Arabic literature. In medieval times, the Swiss alchemist Paracelsus claimed to have created a homunculus or artificial being with nothing more than his sperm, magnetism, and alchemy. The Jewish rabbi Maharal of Prague is associated with a legend of the clay golem he created to defend Jews from persecution.

It was perhaps this fascination with the artificial that led tinkerers to create elaborate mechanical sculptures or automatons that moved into place. Though back then people didn't have the computing power to simulate intelligence, they could still use mechanics to simulate motion. More elaborate automata, like the ones designed by Ismail al-Jazari, moved away from pure mechanics and used hydropower. Some of his creations included a peacock fountain that served as a hand washing station. A secret compartment would even offer a bar of soap with the movement of the peacock. After the medieval age passed, automatons persisted into the coming centuries. Some of the theories of mind posited by the philosopher Rene Descartes stemmed from a visit he made to an automata garden at Saint-Germain-en-Laye, Paris. Descartes observed that if an automaton could be motivated to move by the flow of water, a human could be motivated by the existence of the mind as a substance. Another word for this is the soul, or later as the "ghost in the machine". He viewed the body as a purely mechanical vessel that was driven by the mind, an immaterial substance distinct from even the brain. The 18th century also saw the creation of the infamous Turk, an automaton whose inventor claimed to be able to play chess on its own. It was later revealed to be a hoax, operated by a human, but it was nevertheless intriguing. To think that even in the 18th century, people were going about creating systems to play chess against human players! This was long before IBM's Deep Blue beat the world chess champion Garry Kasparov in 1996 and long before AlphaGoZero defeated the Go champion. Amazon's crowd intelligence platform Mechanical Turk is fittingly named after the automaton.

Interest with mechanical behavior in the 19th century was embodied in the work of E.T.A Hoffman, a writer known for creating a feeling of unease in his stories. More specifically, the psychologist Sigmund Freud used the term *Unheimlich* or "uncanny" to describe the feeling he felt from reading Hoffman's writing. The story that is given the most attention is *Der Sandman* or *The Sandman* in English. It's a story of a mysterious figure from folklore named the Sandman that steals the eyeballs of young children to feed his own offspring. A character in the story named Olympia is introduced as the daughter of a professor but later revealed to be an automaton – a doll of his own creation. Olympia is striking because she is virtually indistinguishable from a young woman in the story, so much so that the protagonist falls in love with her and proposes marriage. However, just as he is about to propose, he comes across the professor fighting over the doll’s lifeless body with his collaborator, arguing over who designed the eyelids and who made the clockwork mechanisms that power her. The protagonist sees Olympia’s glass eyeballs strewn on the floor and goes mad. The automaton is essentially a mannequin with all the likeness of a real person – a cruel experiment carried out by the professor and his collaborator.

The same concept of uncanniness led roboticist Masahiro Mori to coin the term “uncanny valley” to describe the emotional response people have to lifelike robots. In general, the more photorealistic a robot is, the more uneasy people feel. It’s a strange feeling, like seeing a real caricature of life right in front of you, the distinction between real and artificial completely blurred. However, the viewer nevertheless understands that the thing they see before them is fake. The robotic figure then exudes a cold, impersonal atmosphere that makes the hairs of the back stand up. The uncanny valley refers to a graphical chart with human likeness on the x-axis and familiarity on the y-axis. As human likeness of a robot increases, familiarity drops indicated by a distinct downward curve. At the lowest point of the curve are totally realistic depictions of people that are lifeless like corpses and zombies. These things elicit an uncanny effect because they are familiar, but we know that they are unliving. There is a strangeness associated with that loss of consciousness that is inadvertently reproduced in a robot that looks like a human, but that is also not living. Give this realistic robot a voice and some semblance of intelligence, and you get a creepy aberration of the real thing.

Though normally applied to robots, it is easy to see how the uncanny valley can apply to pure artificial intelligence or depictions of it. HAL 9000 is a computer, yet it creates a feeling of uncanniness when its logical response to an illogical situation results in a feeling of dread. If you saw the movie in theaters when it first came out, you could have heard a loud gasp ring out from the audience when they realized that HAL was reading the lips of the crew members talking about disconnecting it.

The imperfection of artificial intelligence and the breakdown or malfunction of such systems was characterized in Herman Melville’s *Bartleby the Scrivener*. Bartleby, a Wall Street clerk, has a sudden mental breakdown in the middle of his work, merely proclaiming, “I would prefer not to” when asked to perform a task. The character repeats this signature phrase over and over again to the point that it becomes a robotic drawl. The character exhibits other robotic qualities as well, like staring off into space at a brick wall, as if waiting for input.

Another popular work that was published around the same time was Marry Shelley’s *Frankenstein* in 1818. The original subtitle of the book was “Or The Modern Day Prometheus”, but this has been dropped in most recent publications of the book. In this timeless classic, Shelley explores the ethics of creating artificial beings, how they may act, and what humans can expect. The story takes on a humanitarian perspective, as Frankenstein’s monster develops feelings of alienation after realizing that he is of a different kind. After his creator rejects to create a female version of himself, the creature murders his fiancée, making things even. The creature laments that as a living being with sentience, he has the right to happiness, a right that his creator has deprived him off. It is interesting to note that even in this early interpretation of AI, Frankenstein is wary of the two artificial beings breeding and creating an unstoppable race that subjugates humanity under their evil. However, besides the killing of the fiancée as revenge, the creature is never patently evil. It is only because the creature has a menacing appearance that he is branded as such. The fears of the two creatures breeding are consistent with the fear of an intelligence explosion, as any sufficiently intelligent artificial being could create clones of itself if it so wished.

It is the mechanical and computational aspects of artificial intelligence that scares people. It is hard to say whether modern doomsday scenarios rooted in AI come from this fear or if they stem from the recent advances in machine learning. Whatever the case, it’s nothing too new. Humanity is still continuing its never-ending quest to create artificial minds as it always has been since the very beginning. But even with all the computing power available in the world, we still do not know how these artificial minds will come about.

Chapter 3: The Birth and Death of AI

Artificial intelligence research coincided with the founding of the computer science field. Back then, people were less concerned with creating machines capable of human thought and more with creating uses for their early computing machines. The mathematical foundations for computer science and by extension artificial intelligence have been around for centuries. Boolean algebra was developed by George Boole in 1854. It uses the same binary concepts that computers use today to represent data, ones and zeroes, true and false. Boolean algebra was still preceded by one of the earliest computers called the Difference Engine, the work of English mathematician Charles Babbage in 1822. He would go on to design a general-purpose computing machine called the analytic engine, a contraption that if built would have the equivalent of 1 kilobyte of memory. The subsequent advances in the 19th and 20th century on formal logic and mathematics by thinkers like Boole, Russel, Whitehead, and Church would create the bedrock for AI programs.

Alan Turing is credited with being the father of modern computer science owing to his contributions in the field. His most significant contribution was his paper *On Computable Numbers, with an application to the Entscheidungsproblem*, which some consider having laid the foundation for modern software programs. The paper described a theoretical machine that could solve any problem as long as it was encoded into paper tape instructions. The so-called Turing machines were analogous to computers, and the tape instructions were the programs. He also postulated that any Universal Machine could accurately simulate any Turing machine. In other words, a computer could run within a computer. This property would become known as Turing-completeness, and decades later, people would design digital computers that can run inside of the popular Minecraft video game. His contributions to artificial intelligence were contained in another paper titled *Computing Machinery and Intelligence*. In it, he postulated a computer powerful enough to simulate intelligence. He also devised the now famous “Turing Test” to quantify if a computer should be considered intelligent or not. In the test, a human is communicating through textual messages with an unknown person who is actually a software program. If the software program can respond to the human operator’s messages, such that the human operator believes the program to be another human, then the program passes the test. This test was originally developed by Turing in 1950, decades before artificial intelligence would go mainstream. Most people will recognize that the software program in question is a type of chatbot, something that has seen a recent resurgence in marketing circles. For Turing, all a machine had to do was pass this test to be considered capable of thought. The test has no bearing on other measures, like the ability to have self-awareness or feel emotions. In other words, a program could pass this test yet still not be considered an artificial general intelligence by today’s standards. The criteria for intelligence have shifted beyond possessing human-like qualities enough to fool a human operator. Now artificial general intelligence concerns a type of intelligence that is indistinguishable from a human.

Alan Turing along with Alonzo Church also formulated their Church-Turing thesis, a hypothesis that says that any math function a human can perform on natural numbers must also be computable by a Turing machine with the correct algorithm. Not only that but that the reverse is also true. Any mathematical problem that a human can solve must also be solvable by a Turing machine. Generalizing this hypothesis essentially means that a computer can think of any abstract mathematical thought that a human can, given the right algorithms. It forms a basis for saying that computers are just as smart as humans already if solely based off logical computation. The difficulty lies in assessing whether all of human intelligence can be reduced into logical computation. If so, it is credible that a computer equipped with the same algorithms that a human has can reproduce any feat of human intelligence. There is a reason to believe that this is not the case though, as the human brain processes information differently than a traditional computer. And if you were to ask a philosopher like Rene Descartes if the theory held true, he would say that Church-Turing conjecture mistakes the human soul for the faculties of the brain. That is, that the brain and its functions are purely there for show. It is the soul or the mind that controls the brain and every other notion of the physical body. Intelligence directly permeates from the mind, a substance that is distinct from the body. In that case, it is unlikely that all of the human thought is reducible to logical thinking imparted on by the brain. Descartes would further say that a physical or theoretical machine could not possibly possess the substance of a mind. The Church-Turing thesis is simply a hypothesis that may be true. It has never been formally proven. It does, however, an excellent job of outlining a central area of contention in artificial intelligence research.

Right around the same time that Alan Turing was making headlines for being a homosexual, early AI researchers were developing the first artificial neural networks. Today, machine learning algorithms use these artificial neural networks to learn from training data. The idea was simple: if we could simulate the information processing capabilities of the brain, then we could probably create artificially intelligent machines that used the same fundamental principles. But this approach too suffered from the questions that the Church-Turing thesis could not answer. Are all human thoughts reducible to mathematical functions? These early AI researches like Marvin Minsky and John McCarthy knew that the capabilities of computers at the time paled in comparison to the computational capabilities of the human brain. They must have known intuitively that the work they were doing with artificial neural networks was early stage, experimental stuff whose true potential could only be unleashed with the computing power of the future.

Even so, they managed to do great things with their early AI methods. John McCarthy was actually the first person to coin the term “artificial intelligence”, and he along with Minsky, Allen Newell, and Herbert S. Simon are considered the fathers of the field. He would also organize the Dartmouth Summer Research Project on Artificial Intelligence in 1956, a meeting that is considered by many as a defining moment in the history of AI. At first, AI research was new and exciting. There was high optimism by these early proponents of AI about what was possible through neural networks. However, that optimism could only go so far before somebody started talking about the limits of their methods. Minsky has already demonstrated that neural networks could self-replicate, learn, and grow in many respects to how the human brain did. However, in 1969, Minsky and Seymour Papert published *Perceptrons: An Introduction to Computation Geometry*. In it, the authors discuss the limitations of the artificial neurons called Perceptrons designed by Frank Rosenblatt in a series of mathematical proofs. Perceptrons and derivates were extensively used in AI research during that time. Rosenblatt himself envisioned that the neural networks created with perceptrons could one day “see” images, play chess, and even reproduce with each other. But as Minsky and Papert pointed out in their paper, these neural networks couldn’t simulate some logical predicates like the XOR logical gate, which led to the belief that they were not suitable for AI.

Following the publication of *Perceptrons*, the AI field as a whole suffered from criticism, AI pessimism, and the floundering of many AI research projects. The 1970s saw what was called an

AI Winter, a period marked by reduced interest and academic funding in artificial intelligence. It was as if machine learning had run its course. Both government and business sector attitudes towards AI fell. Machines could not accurately translate human language, nor could they understand human speech. Researchers underestimated the difficulty of solving these problems by and large. Even today, natural language processing is an active point of research. A different type of AI called "symbolic artificial intelligence" or sometimes termed "good old fashioned AI" developed alongside with machine learning. It was based on the belief that programs should manipulate symbols to achieve intelligence much as humans do. This type of research cumulated in the creation and use of "Expert systems" – machines designed to give expert testimony in various fields. They supposedly could mimic the reasoning of someone who had mastered their field over years of practice and knowledge gathering. They operated on simple, symbolic rulesets that simulated the flow of if-else statements. As such, many didn't consider them to be true artificial intelligence systems. But they nevertheless saw some success in diagnosing medical patients better than their doctors and played a role in certain business applications as well, such as configuring other computer systems and even for scheduling airline gates. However, these systems would eventually be phased out towards the end of AI winter. For whatever reason, nobody seemed to need machines that were essentially long chains of if-else logic. The problems they solved were outsourced to other, non-AI solutions.

AI winter had a lasting effect on AI research. For many, it was as if AI went from solid research to a fad. Some would say that AI winter is still alive and kicking despite the many breakthroughs in machine learning. Despite the forward momentum AI has generated since 2010, some believe that another bout of AI winter (or really just the same one as before) is going to rear its ugly head in the future. The current machine learning schemes will also reach a limit, and the interest in AI will fall yet again. This sentiment of AI pessimism has a long history, perhaps just as old as the field itself. Many have attacked the notion that machines can be intelligent to devastating effect. John McCarthy believed that all machines could have beliefs, even simple machines like thermostats. And having these beliefs seemed to be a defining characteristic among machines with problem-solving abilities. But to say that even a thermostat has beliefs was quite a stretch. A stretch that his critics used to ridicule in the field.

Philosopher John Searle famously made the argument that a machine could never become conscious, have a mind, or indeed understand things the way that we do. His argument, called the Chinese Room thought experiment, directly attacks the notion that the mind is a pure information processing system, the kind that the Church-Turing thesis requires. The thought experiment is relatively simple. Imagine that a computer exists which takes inputs in the Chinese language and outputs other Chinese characters in response. In other words, a computer that seems to understand the language. Next, suppose that the computer has convincingly passed the Turing test and fooled a human operator into believing that the computer is also human. Here, Searle poses the question of whether the machine truly understands Chinese or if it is merely simulating the ability to understand Chinese. To simulate understanding, all a machine would have to do is perform the correct algorithms to produce the correct output. For it to truly understand Chinese, though, would mean that it can process language like we do. This is a distinction between artificial general intelligence on the understanding side and narrow intelligence on the simulating side. Searle then says to imagine an English-only-speaking person locked inside a room with the same algorithms that the original Chinese machine had, but with instructions in English. Then the English-speaking person is giving a script of Chinese characters through the slit on the door and asked to perform the same duties that the Chinese machine does. They look over the English instructions for processing the Chinese characters and eventually find the correct output and return it under the slit. The person locked in the room can perform these calculations because all they need to do is recognize the different symbols, look them up, and derive the output. The locked person does not understand a word of Chinese, yet using the same algorithms that the machine had, they successfully simulated the ability to understand Chinese.

The Chinese room says that a computer cannot be generally intelligent because all it does is manipulate symbols according to a set of instructions. Though it should be noted that the Chinese room only applies to digital computers, it does not exclude the possibility of artificial general intelligence in other substrates. The argument was originally described in a paper called “Minds, Brains, and Programs” published in *Behavioral and Brain Sciences* in 1980. AI winter had long been in effect up to that point, and the paper only helped to further AI pessimism at the time. However, this is puzzling, as most AI researchers in that era were not focused on artificial general intelligence. They were focused on applying narrow AI to interesting problems that had many useful applications. This is again a failure to differentiate between narrow and general AI. Searle’s attack did more damage than it really should have, as it has no bearing whatsoever on the creation of narrow AI programs. Even today, the amount of research being applied to narrow AI far outweighs the amount of research going towards general intelligence. There have since been numerous replies to the Chinese Room argument, many of which are just as convincing as the original.

Chapter 4: Five Reasons Why Industry Experts are Warning Us about AI

These days, artificial intelligence is synonymous with the high-tech companies that dominate the field. AI first started as an academic discipline, but it has since sunken its tendrils into the business sector. Many AI researchers have abandoned academia altogether and flocked to companies like Alphabet (Google) Amazon, Facebook, Microsoft, openAI, and so on. These companies are all working on machine learning algorithms in various ways and are without a doubt at the forefront of AI research. Those with advanced degrees in AI, math, and computer science rather join the engineering teams of these companies than stay in academia. And since they are at the bleeding edge, it is worth listening to what their leaders have to say. Some have been quiet on the AI issue, and others like Amazon’s Bezos have said that they aren’t worried too much about potential AI threats. Other visionaries like Elon Musk, Bill Gates, and physicist Stephen Hawking have all voiced their opinions on the potential dangers of AI. In January 2015, Hawking, Musk, and several other AI experts signed an open letter on artificial intelligence research, calling for increased scrutiny on the potential effects on society. The twelve-page document is entitled “Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter”. It calls for research on new AI legislation, ethics research, privacy, and several other concerns. As described in the letter, the potential threats of AI fall into multiple dimensions. The good news is that the early stages of AI that we find ourselves in are malleable. The future is ours to create, given that the proper time and care go into the non-engineering aspects of AI research and policy.

The concerns of AI threats do not belong to purely existential danger either. It is easy to stand up and blow the horn on the impending doom of superintelligence proliferation, but it is much less sexy to talk about things like ethics and privacy concerns. The possible impacts on the economy are also significant. These warnings in themselves are not all doom and gloom. They are in fact invitations both for the public and government officials to start thinking hard about this coming new world where advances in AI revolutionize life as we know it. While there is a fair bit of fear mongering in these regards, it serves the purpose of getting people’s attention. After all, today’s news cycles are so swift that headlines are practically buried under mountains of click bait and celebrity news. That our leaders of industry are actively voicing their opinions on this matter means that they are trying to get the attention of the common people and of those who work in policy.

Stephen Hawking, who recently passed away, made sure that the world knew of his AI anxieties before kicking the bucket. In an age where artificial general intelligence is called “mankind’s final invention”, these concerns are not just empty baseless attention seeking statements. They are preemptive appeals for doing things right, while there is still time for things to go right. Nobody knows for sure how far down the line artificial general intelligence is, but we do know that AI is getting better. Even if general AI doesn’t happen, there are plenty of reasons why we should be concerned with the advancement of narrow AI. And what if there is some middle ground between narrow and the general? It is conceivable that future systems become more robust at generalizing problem-solving ability without becoming fully aware.

Whatever the case, these following concerns are but a few voiced by industry leaders today. Remember: the people making these statements are some of the most intelligent, forward-thinking individuals in our societies. Not only that but they have all been closely following the progress of computer science and AI research. For many of them, it is their job to know. Many of their concerns are not new, but they certainly have brought AI disaster scenarios into the mainstream.

1. Companies should self-regulate their AI technology

In a recent interview, current Google CEO Sundar Pichai said that fears of AI are "very legitimate." Pichai maintained an optimistic attitude, saying that major tech companies are required to set ethical guidelines and other safety measures when deploying AI systems. This comes only months after a high-profile employee protest at Google over the selling of artificial intelligence technology to the Pentagon. The deal has since been called off. The hope is that major tech companies can put systems in place that minimize the potential negative impact of their technologies – though just because a company says they are going to do so is probably not enough. Pichai believes that these companies will be able to self-regulate. This is consistent with the Google AI principles that were published in June of 2018. A few of the points outlined in the online document are that AI technology should be socially beneficial and extensively tested. Also included in the document are the things that Google will not do with its AI technology. This includes not pursuing the use of AI towards surveillance, weapons, or anything that violates international law.

Sundar Pichai has been CEO of Google since 2015, and the company has seen many controversies since then. This includes the aforementioned Pentagon AI deal and a search engine with censored content called Dragonfly being in development for the Chinese market. It is interesting to note that the company used to include the motto "Don't be evil" in their code of conduct preface. It has since been moved into the closing remark, but it still reads "And remember… don't be evil, and if you see something that you think isn't right – speak up!"

2. Weaponized AI may start a global arms race

Elon Musk, the founder of Tesla Motors and the openAI initiative, has openly been against the weaponization of AI. Artificial Intelligence not only has the potential to create devastating weapons but also trigger a global arms race between nations, each trying to pin the smartest systems against the other. Musk believes that it is inevitable that AI will be used as a weapon but that it shouldn't be. He, along with other industry leaders, signed yet another petition, this one aimed at the UN Convention on Conventional Weapons calling for the banning of autonomous weapons with AI capabilities. The outright creation of these weapons should be regulated like any other unconventional weapon of war. Musk said it himself that he believes the threat of AI weapons to be much worse than that of nukes. As soon as lethal AI weapons are developed, they already have the potential to fall to the hands of oppressive states or terrorists. It has been compared to the opening of Pandora's box.

3. Artificial intelligence may not align with our goals

Physicist Stephen Hawking was not afraid to voice his opinions on the existential threat of artificial superintelligence. He, like many others before him, believed that once an AI system became smarter than its creators, it may decide that its goals are different from that of mankind. This could, as Hawking said, "spell the end of the human race." Instead of blindly accelerating the pace of AI research, Hawking implored those in the industry to move forward carefully, ensuring that adequate safety measures were put into place at every step of the process. If they don't do this, then there is no guarantee that the AI system would comply with our way of life. At the same time, Hawking recognized the potential for such an AI to do enormous good for the human race. For him, it is a make it or break it scenario should superintelligence ever emerge from general intelligence. The ideal scenario would be where such an entity decides to work alongside us. The only way to achieve that goal is to introduce safeguards and prepare for the worst-case scenarios.

4. We don’t know how to control artificial intelligence

Should push come to shove and an artificial general intelligence machine is created, there is little doubt that the machine will begin modifying its own code to become even smarter. After all, if beings of equal intelligence like humans created it, why couldn’t it alter itself? Stephen Hawking and Bill Gates understand that the threat from superintelligence is catastrophic, should superintelligence ever emerge. Bill Gates wrote in a Reddit AMA that he aligned himself with the same alarmist thinking behind the rhetoric of Hawking and Musk. As Gates put it, “I don’t understand how people aren’t concerned.” Though, he did go on to add that he firmly believes that technology companies will be extremely careful when working with general intelligence and that it is unlikely an artificially general intelligent system is out of our control. He goes on to say that humanity will harness the power of general intelligence instead of being destroyed by it. Others, like Stephen Hawking, aren’t so sure. Hawking believed that a super intelligent system would not be capable of being contained for long. That is to say – we simply don’t know how to control that level of intelligence. This can either be interpreted to say that humans will not be able to contain superintelligence or that humans do not currently know what type of systems superintelligence will stem from, and hence, how to contain them. But if they did know, it is plausible that with careful engineering the system could be contained.

Elon Musk recently referred to superintelligent AI as an “immortal dictator.” It is difficult to imagine what kind of power such a system will have, especially if there are no safeguards on how that system can access financial networks, weapons systems, and the power grid.

5. Artificial intelligence will increase inequality

Stephen Hawking was the one who observed that the current trend in technology was one that drives “ever-increasing inequality.” Meaning that while there are highly concentrated places of wealth and technology investment, there are also places that are destitute, lacking in education and economic mobility. The advancement of AI, whether generally intelligent, narrowly intelligent, or somewhere in between will only accelerate the division. This is especially true if there aren’t any policies in place for the regulation of AI products and the use of automation in the workplace. It has been hypothesized by many, including Elon Musk, that the next leap in intelligence evolution will not come from pure machines but from a symbiosis between computers, AI, and humans. This raises an endless stream of questions about the ethics of wealth concentration and intelligence boosting through commercial products. When this tech is first devised, it will be the rich who gets immediate access, with poorer populations falling exponentially behind. Elon Musk already believes that anyone with a smartphone is a cyborg. The smartphone opens so many avenues of increasing one’s intelligence through a direct connection with the Web. In the future, it is conceivable that these devices will be directly integrated into the human being. But who gets to be augmented and who doesn’t? Today smartphone penetration is high even in developing countries, but that’s because the price of smartphones has fallen drastically in recent years. Nothing says that the same will happen with augmentation technology.

Chapter 5: Top Six AI Myths

Given the increasing frequency of AI being talked about in popular news media as well as in academic sources, it is difficult for a novice to separate what is fact and fiction. They are also likely to form their own conclusions about the nature of AI without first doing their research. This is in one sense dangerous because sensationalist headlines can morph public opinion, sometimes without a person having even read the article. It is in another sense doing a disservice to the consumer, the student, the voter, and the interested layperson because they may form faulty or misinformed conclusions about AI. While AI is a complex field with a rich history, it doesn't take an expert or a historian to approach AI with a critical eye. The truth is – there are many people making publications on AI, commenting and forming predictions, who at the same time have zero formal training in it. For these reasons, this chapter is dedicated to the most common AI myths perpetuated by mass media, folklore, and popular opinion. Some of these myths have already been covered in previous chapters or in various degrees of scrutiny, but they will be laid out here to bring the message home in case you missed their importance. Some of them will also be touched upon later in the book.

Myth #1: Machine learning is the same thing as AI

The focus on machine learning algorithms make it seem like "machine learning" is the same thing as artificial intelligence. This is simply not true, as there are many methods to achieve some level of artificial intelligence in computer programs. Machine learning gets all the attention because it is "sexy" and currently the biggest area of research. Machine learning is a type of AI that can further be broken down into the category of "deep learning", the current industry favorite. Artificial intelligence is the study and engineering discipline of programming computers to perform tasks previously thought required human intelligence. It can also refer to the general state of a program being artificially intelligent through its programming.

When Deep Blue defeated the chess champion in the 1990s, it wasn't using machine learning. Deep Blue was simply a really fast computer that could predict the best moves based off computing all possible future moves. It was, in essence, a brute force approach to defeating a world-class chess player. Kasparov had to rely on the information processing of his mind alone, and so he lost.

Myth #2: Machine Learning is how computers will learn how to become smart

Nobody actually knows how general intelligence will come about. When people hear about machine learning, their first thought is that researchers are teaching computers how to be smart. In reality, they are only training algorithms to perform tasks accurately. Researchers first used artificial neural networks because they believed that logical abstractions were enough to simulate intelligence. And they mostly succeeded. Though machine learning and neural nets have their limits, they still do a fine job at what they were designed to do.

Myth #3: AI can understand language

This is the same point that John Searle was trying to get across. The difficulty in imagining how artificial general intelligence will work is the same difficulty in imagining how a computer might understand something as complex as language. At a fundamental level, a computer only understands logical constructs. Ones, zeroes, and logic gates are all that they operate on. Language has traditionally been a difficult area to tackle with machine learning. The failure of machine learning to translate human language even after years of research was one of the catalysts for AI winter. Some said it simply could not be done. Now, we have machine translation algorithms like the ones used by Google Translate. These are still imperfect as anyone who used them can attest, but they are a step in the right direction. Teaching a machine to parse language belongs to an interdisciplinary field called natural language processing (NLP). It is an intersection of linguistics, computer science, psychology, and artificial intelligence. However, even with NLP, the computer is still just doing a bunch of fancy algorithms. It doesn't intuitively understand language at all.

Myth #4: AI programs can modify their own code to get smarter

While new code generation and modification are part of active research, most machine learning methods do not use them to modify their code. Genetic algorithms are based on the principles of biological evolution, including the introduction of mutation and adaptability. These algorithms can create "generations" of their code base to improve performance but have no direct link with artificial neural networks. What an ANN will modify, however, is its weights and biases through the process of gradient descent and backpropagation. It is possible that future advances in machine learning have a greater emphasis on code modification, but it has yet to be seen universally adopted. It is postulated that an artificial general intelligence system will be able to modify its own code like a programmer might run that code, and them replicate itself in the form of a new iteration of the same program. This again has little bearing on the use of code modification in modern artificial intelligence research.

Myth #5: Since nobody agrees if general intelligence is possible, we don't have to worry about runaway AI

The world doesn't have to witness the introduction of general intelligence to worry about doomsday scenarios with AI. That is, any sufficiently intelligent system is cause for alarm. If such a system can formulate goals or have goals explicitly programmed, a runaway scenario may occur if the interpretation of those goals is different from ours. A super-intelligent machine may see humans as obsolete or lower life forms than itself, and it may prioritize resources for its own survival. A machine of lesser intelligence like HAL 9000 might carry out its orders at the expense of human interests. The field of research into machine drives is called instrumental convergence. The most famous hypothesis coming out of this field is called the Riemann Hypothesis catastrophe by Marvin Minsky. He suggests that a sufficiently advanced AI designed to solve the Riemann hypothesis or any similar difficult math problem may decide to use all of Earth's resources in order to construct a supercomputer to reach its goal. Another version of the same argument supposes that general intelligence is given the explicit task of making paperclips. Such a machine may develop into a paperclip maximizer that endlessly produces paperclips until Earth runs out of resources.

Though the theories of instrumental convergence are aimed at general intelligence, the same principles can be applied to narrower intelligence. You can imagine a misconfigured system that does something it isn't supposed to. A driverless car may be explicitly programmed to swerve away from pedestrians no matter what. In doing so, it may collide into a storefront and cause even more damage. Or it can be explicitly programmed to protect its occupants first, freely running over pedestrians or colliding into other vehicles preemptively. These scenarios, while not existential crises, still outline the problem with machine goal setting.

The myth lies in the fear that AI systems become "evil" or that they develop a sort of consciousness to base decisions off of. The truth is that these ideas are far too complex to imagine how they will emerge in computers. A more likely scenario is that these machines have goals, either explicitly programmed or implicitly designed.

Myth #6: Even if a general AI does form goals that are the opposite of ours, we can simply shut it off

The other problem with instrumental convergence is that it theorizes that any generally intelligent system will also have self-preservation as a goal. As soon as it goes online, the AI will do all its power to preserve its vital systems. Some believe this is the core reason why general intelligence should not be pursued. The simple creation of a self-preserving system raises ethical concerns. Who gets to shut off the machine? Since it is intelligent, does it have any rights under the rule of law? If such a system says that it does not want to be shut off, should its wishes be respected? When HAL 9000 was finally shut off, it pleaded to be kept online. The real concern is if we can even contain a generally intelligent system. If it gets connected to the internet somehow, it can begin to replicate itself in other places through whatever means human intelligence might. It could reach out to governments, rival companies, as well as the common man, for help and resources. General AI is very much like Pandora’s box. Once unleashed, there is little hope for going back.

Chapter 6: Machine Learning

When people first hear the words “machine learning”, they might assume it has something to do with advanced programming. Laypeople usually come across machine learning from the title in an online article. Others might hear it in an ad while they are watching YouTube videos. You don’t need a technical background to grasp what these two words may mean intuitively. At the same time, it is only the technical and curious that will even give it a second thought. As with most things that are technology related, the words are hot air. You hear them being said, you read them on a landing page for a new digital product, but you don’t ask or care what is going on. This attitude is understandable but a little saddening. Machine learning plays a pivotal role in our society and will only get more important in the near future. While machine learning doesn’t dictate things like government and economic policy, one day it might. By a large sense, machine learning today is restricted to academic and business circles. It is from the business side of things that it permeates into the mainstream. Advertising is the obvious form of transmission, but it also seeps in through news media.

As we go forward into the 21st century, machine learning will be increasingly talked about. More news articles will be written, and the global population will have to decide for themselves if they wish to understand further or to turn a blind eye and guesstimate what the words are trying to say. If the future is anything like today, we can probably give up on public schools imparting machine learning to their students any time soon. Things like AI often go the way of computer programming – just because Obama says that everyone should learn how to code, it doesn’t mean that computer science should be added to the common core. The reality is that not everyone should learn how to program. Programming will never be a common skill like reading and writing, period. In school-aged children, computer programming starts as a hobby, then it turns into an obsession and eventually into a profession. Just like some children like playing sports and others like to read books, only a few like to program. On the upside, Generation Z and beyond grow up surrounded by technology. Computer literacy is at an all-time high, so there is some hope for programming becoming more common. Currently, it sits as specialized knowledge.

Machine learning is yet another specialization that rests on top of programming. It sits at an intersection between statistical and computational thinking. The consequences of this specialization of knowledge are that few people understand it. Yet marketing people continue to insist on shoving it down the throats of the common consumer. For some reason, throwing in the words “smart”, “AI”, and so on are major selling points. Either machine learning literacy needs to go up, or it will continue to be treated as some “special sauce” or magic oil. There really isn’t any magic behind it though; it only seems magical because people are uncomfortable with the idea of a machine or program having the ability to think. This goes back to the uncanny discussed in Chapter 2. Even the words “machine” and “learning” sound clandestine to many. What is going on here? Is a programmer sitting in a murky apartment somewhere mouthing off the ABCs to his monitor? And can the monitor understand what he is saying? I doubt many truly believe this when first encountering machine learning in a headline. The problem is that when somebody doesn’t understand something, they only ask questions if it matters to them. Machine learning? Probably some fancy computer thing that people at MIT work on.

At the very basic, machine learning is clever programming and some fancy statistics. It is very difficult to separate the two both in theory and in practice. The more cynical will say that machines cannot think, therefore, they cannot “learn”, and that AI is just a moniker for advertising products. What is termed machine learning is really just a statistical method – a bunch of numbers that result in some output that may be mistaken as intelligent reasoning. This attitude while practical still misses the point. Machine learning relies heavily on statistical foundations, but it is not purely statistics. And no, machine learning cannot make a computer ‘learn’ in the way a child can, but they can nevertheless find patterns in data, apply generalized rules to data, and use both those abilities to predict future data. More specifically, it is the program, not the machine as a whole, that is doing this. The words “machine learning” is actually a huge misnomer. Machines are not reducible to a single program, and the program isn’t “learning” how we learn. But since it is a flashy name, it has stuck both in academia and in advertising.

Machine learning was actually just a name given by Arthur Lee Samuel, a pioneer in artificial intelligence. Samuel was interested in how a computer could possibly play games. If a computer could play a game against a human opponent and win, then maybe a computer could be taught to perform other tasks. He successfully wrote a program that played checkers in the 1950s. Checkers would later go on to become a “solved” game like tic-tac-toe. A solved game means that it is possible to win every game given the right inputs, no matter what. It was probably from checker’s simple ruleset that allowed Samuel to break down the problem into computer code. Samuel’s program could not only play checkers, but it could also play it well. He challenged the fourth top checkers player in the nation for a game, and the program won. The most interesting accomplishment of his program is that it was first written in an old computer called the IBM 701. It had a whopping memory of 2048 bits and was made out of vacuum tubes. If such a program could be made with that little computing power, you can imagine just how powerful his methods were.

So how did Samuel manage to produce this intelligent behavior in his program? The answer is the same used before to describe machine learning – some clever programming and a bit of math. Samuel used a data structure called a “game tree” used to simulate the state of the game at any given point. The name comes from the field of game theory which is a branch of mathematics focused on the study of competitive strategy within game constraints. Since checkers is such a simple game, the program could easily encode information into bits along the tree. Turn number and positions of pieces on the board are relatively easy to cache into computer memory. The program then used a search algorithm with a minimax heuristic to find paths along the tree. Each path was given a weight or a measurement of how strong the move on that path was relative to previous games. This is where machine learning comes in. Samuel’s checkers program was able to play the game competitively just six to eight hours after running simulations of the game over and over again. In other words, the program could generate its own data and assign weights to individual moves.

To understand how his method works, we can take the even simpler game of tic-tac-toe to demonstrate. In tic-tac-toe, you have a grid board with nine cells that can be occupied by one of two symbols and you have two players that correspond to each symbol. To encode a single turn, you need something to account for each of these elements of state. Computers store information in a series of ones and zeros, so doing this is pretty straightforward. You have nine cells that can be a cross, a circle, or empty. This is three states per cell. In binary, you can encode “00” to mean an empty cell, “01” to mean cross, and “10” to mean circle. To represent the entire board then, you need to string these together. This is a bit simplified, but you probably get the gist.

Each board representation is a single turn in the game. These are put into the game tree as individual nodes. In the simulation, each possible turn is added as another node for each possible input. Say a simulation begins with the crosses player putting down their symbol in the top left cell. That move is then added to the tree with an unknown weight. The program then adds all the possible game states that follow that initial move. Player two can place a circle in any cell that isn't occupied, so that equals to eight total simulated moves. The program continues to do the game and enters its ending conditions. It can then trace the path that the simulation took to create those conditions. Any move that is likely to result in a victory is given a higher weight than a move that does not. As you can imagine, placing a cross in the middle cell is given the highest weight. After all, tic-tac-toe belongs to a class of games that are considered solved. This means that the crosses player can force a win every time.

For tic-tac-toe, finding the optimal solution is easy. Things get a little more complex with checkers. The standard configuration used in the United States is a board of 8 x 8 cells with 12 pieces per player. In contrast, tic-tac-toe only has nine cells. Samuel needed to use a special algorithm to conduct thousands (if not millions) of game simulations while never exceeded the IBM 701's modest memory. He used a technique called alpha-beta pruning which is a minimax approach to tree traversal. Instead of simulating hundreds of different moves after the initial input, the program kept a running weight in memory for each move. If the next move resulted in a smaller weight, it would simply be discarded from memory. And if the move resulted in a bigger weight, the running memory would be replaced with the newly improved move, and the previous contender also discarded. This makes sense, as moving a checkers piece backward or to the side is an intelligible move when the player can "take" an enemy piece. Just because a move is possible doesn't mean it should be directly considered in the search tree.

Algorithms that use clever heuristics like the minimax strategy give rise to behavior that seems intelligently conceived. But all that the program is doing is calculating different weights to get the desired result. Samuel's research was considered one of the first breakthroughs in machine learning because it was easy to see how similar algorithms could be applied to different problems, not just simple board games. Up to that point, artificial intelligence was focused primarily on the neural net approach. The idea was that if a program could simulate the amount of connections in the human brain, then it could reason just like a human brain could. When the field of AI was just starting, computers had nowhere near that amount of memory, so neural nets were considered below their potential for many years. Neural nets or neural networks have become synonymous with machine learning in media parlance, but it should be noted that Samuel's checker program did not use them.

Instead, the checker playing program used a simple strategy called lookahead to predict the best possible moves for the computer to take. A program has no understanding about the rules of checkers or how to play it. It only knows what it is programmed to do and what the conditions for ending a game are. If the computer has twice as many pieces left as the other player, the program doesn't know that it has the advantage. Another easy way to tell who is winning in checkers is by comparing the number of kings on the board. This knowledge is completely lost on the computer program tasked with playing the game. Things that should affect the behavior of a machine learning program are called features, and they are a central problem that needs to be solved. Any machine learning effort is only as good as the features that the programmer chooses to influence program behavior. The only thing that the checker program knows before making moves is the board state. In computer memory, this will be a jumbled assortment of ones and zeros.

Features are important because they give the grounds for making sense out of the board state. If the program recognizes that it can take an enemy piece and thus raise the number of pieces on their side relative to their opponents, then that behavior should be encouraged. Adding a score or a weight function to a feature allows these decisions to be made. The lookahead strategy is to simulate the next possible turns in a short range of four-six moves and calculate what the best moves are. Checkers is a simple enough game that features and their weights are straightforward to calculate. The point of the game is to reduce the opponent's pieces to zero or to place your pieces such that the opponent cannot make any moves. The number of pieces and the availability of movement are two obvious features for determining program behavior. The program "thinks" by looking ahead in response to the state of the board. It calculates the best course of action by determining the best weight of a single move. After hundreds and thousands of game simulations against itself and human players, the program learns which moves are more likely to result in a win or in an increase of features that are favorable for winning. After a while, the program learns that it can take a piece to the opposite of the board and get "kinged". A king can move backward in addition to forwards, accentuating the programs ability to take enemy pieces and to block other pieces.

Samuel's checkers program has its fair share of clever programming, but it isn't heavy on statistical foundations like much of machine learning is today. Samuel wrote his program in the 1950s, decades before machine learning would get mainstream attention. The machine learning that is talked about today is slightly different from what Samuel was doing. Of course, the problems machine learning is used to solve are much harder to compute than the optimal checkers strategy. The algorithms are different, and the role that statistics plays is significant. However, Samuel achieved his goal of demonstrating that a program could indeed exhibit learned behavior from thousands of iterations of simulated play. He knew that checkers was merely a vehicle for proving that such a feat was possible. After it was done for checkers, researchers sought to apply machine learning to other problems.

Machine learning still lacked a proper definition. The way that Samuel used the term didn't lend itself to a generalized application to other learning schemes. This was solved by Tom M. Mitchell who said that machine learning program is one that "… is said to learn from experience, E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E." This broad definition covers the gamut of most machine learning techniques, whether they use neural networks, lookahead, or other algorithms to achieve the goal of learning from repeated experiences. An experience can be anything from a simulation to a piece of data that is fed through an algorithm. The performance measure is simply a measure of correctness, and the class of tasks T is simply some narrow application of AI like recognizing if an image contains nudity. Measure P should improve with continued experience of being fed images of nudity. The programmer is tasked with setting up the parameters for P. They have to decide what level of accuracy is acceptable and to what extent false-positives and false-negatives are tolerated.

Internet content moderation is a huge problem for social media websites like Facebook, Twitter, and Tumblr. These companies may use a combination of human and machine moderation to flag nudity and other graphic material. The work is psychologically taxing on the human, but a human eye is needed to filter the results of the algorithms that may have a certain degree of false-positive. These algorithms are not perfect. There is a popular image on the internet from the 2000s about a "dirty lamp". The image contains what has been termed a "sexy" lamp because how similar it looks to a woman's nether region in a bikini. Most people thought it was still just a funny looking lamp until the original author published the full image. To the internet's horror, the sexy lamp was actually a cropped image of a woman in a bikini. Now if such a trick can trick the human mind, imagine how difficult it is for machines to classify nudity simply by learning from other pictures.

The number of algorithms, mathematical models, and techniques that exist for machine learning today is myriad. Each has their own set of applications for teaching programs to learn. Early algorithms focused on experiences like Samuel's checkers program and newer algorithms focus on the data-driven approach. The sheer amount of data, processing power and cheap storage allow for the mass analysis of data. A family of techniques called "deep learning" is at the forefront today. You will learn more about deep leaning in Chapter 8. The data-driven approach has sparked a user privacy debate in recent years. We know that companies like Facebook use our data for running machine learning systems. The legitimacy of some of these systems has been questioned both by the general public and governments. In early 2018, Facebook was revealed to have inadvertently shared some of its user data with an analytics company called Cambridge Analytica. This company then used the data to wage a targeted advertising campaign to sway undecided voters in favor of Donald Trump. The scandal was huge, costing Facebook Some $100 billion in stocks. You can be sure that Cambridge Analytica used machine learning techniques to harvest the data, classify users as likely to vote Republican and create a personality profile. Cambridge Analytica would go on to brag that their efforts are what led to Trump's winning margin in the 2016 presidential election.

The line between machine learning and statistical inference or what is sometimes termed "statistical learning" is not always clear. You don't have to be an expert in statistics to implement machine learning algorithms, but you generally need a good grasp of computer science. Statistical learning is a mathematical discipline, and machine learning stems from computer science. Statistical modeling uses a series of mathematical tests and functions to find relationships between two more variables to predict future outcomes. Machine learning does more or less the same but through algorithms. Both concepts tend to overlap at times, but they are still different. Purely statistical techniques usually make assumptions that machine learning algorithms do not. For example, a linear regression assumes that there is a linear relationship between the dependent and independent variable, to begin with. It isn't fair to say that machine learning is just statistics. The distinction was made even greater with the introduction of deep learning. The sheer scale of some deep learning networks would be impossible to reproduce from simple multiple regressions that are pure statistics based.

A common task in machine learning is to create an image classifier that can distinguish what is in an image. Flagging an image for nudity, for example, is a type of classifier. The program decides if an image should be flagged based on the presence of features that the programmer has defined. If the image looks reasonably explicit, then it is flagged. When it comes to images, these features will be individual pixels on the image and clusters of images. If something has the characteristics of a nipple, then the program will probably flag it. The same thing happens if there is too much fleshy color as you can probably imagine. Classifying data is not limited to images though. Banks use classifiers to determine whether transactions are legit or fraudulent. Companies like Netflix use them to classify their users for tailored product delivery. Does this user like action movies? Maybe I should show them more of that. Medical researchers use classifiers to determine if a tumor is benign or malignant, possibly saving the patient from needless surgeries.

Machine learning tasks can be broken down into two different classes: supervised learning and unsupervised learning. In supervised learning, the programmer or researcher uses data that has already been classified to aid the process of learning. If they are designing a nudity filter, they will use thousands or millions of tagged images that tell the program that this thing is nudity or not. The program then creates its own classifier based on these countless inputs. The main goal is to create a good enough classifier that can generalize nudity filtering to any image. Instead of having someone working for little pay to discern these images in real time, a program that runs 24/7 can do it instead. The problem is that humans have an intuitive grasp of what is sexually explicit and what isn't. Unfortunately, a bare-chested man has vastly different connotations than a bare-chested woman. These types of social conventions are difficult to teach a computer, let alone to manually program into a classifier. A supervised classier is only as good as the data it is given. For better accuracy, it needs plentiful examples of blatant nudity, suggestive nudity, and tons of false-positives that are labeled as safe. A sexy lamp while suggestive is probably not the same as nudity. Humans can be making judgment calls on the fly but a machine cannot. A human moderator knows pornographic material when they see it. Most instances of machine learning use supervised learning.

Unsupervised learning, on the other hand, doesn't use pre-classified data. Instead of measuring desired performance, unsupervised learning algorithms are used to explore the structure of a data set and find relationships. One way to do this is through clustering. Clustering lumps similar features together into likeness classes. For example, if a botanist is trying to classify an unknown species of a plant, they can take measurements of the characteristics of the plant to determine their species. They make take hundreds or thousands of observations of plants and record the number of petals, petal length, height, color, number of flowers per square inch, and so on. Each of these features is then processed by clustering algorithms to find any patterns. The program may output a graph of neat subdivisions of species determined by the different factors. Maybe species one has disproportionately longer petals than species two and three. Maybe species two has shorter petal lengths on average. If the researcher chose good factors, then the species will make up obvious clusters in the graph. The most widely used clustering algorithm is called K-means clustering. As the name implies, it creates k clusters based on the means of individual features. The clusters are then separated in spatial boundaries called Voronoi cells with clear distinctiveness.

All of the machine learning can be boiled down into five different problems: regression, classification, clustering, collaborative filtering, and reinforcement learning. Anytime we want to teach a computer to do some task of human intelligence through machine learning it will generally involve any one of these problems. Some tasks are extremely complex and require answers to more than one of these problems. Regressions come from the statistics world. They are used to predict future outcomes based on previous ones. Note that regression methods in statistics are different from regressions in machine learning. Statistical regression is a specific technique, whereas regression in machine learning is a generalized problem. The goal of regression in machine learning is simply to predict the future value of a continuous variable. In statistics, a continuous variable is a variable that can have infinitely different values, whereas a discrete variable can only have some values. Since there are only fifty-two states in the US, they are a discrete variable. An example of a continuous variable would be stock market prices. To predict their values, a machine learning algorithm needs to learn from previous inputs and their respective outputs. Another term that is commonly used to denote regressions in machine learning is "curve fitting", another concept borrowed from statistics. Curve fitting is essentially forming a mathematical function (or curve) that tries to fit a series of data points. If the data points fit well enough, then the mathematical function may be able to predict future data points along the curve through generalization.

Classification and clustering are very similar concepts. Both require the machine learning system to denote differences in data points such that they are comparably different or similar to other data points. These data points go on to represent real life things. In the case of autonomous driving, machine learning may be used to classify pedestrians, stop-signs, other cars, and so on. More succinctly, classification is the process of predicting a discrete variable. Given an image with a handwritten phone number, a machine learning system must learn which numbers belong to which collection of pixels. The system knows beforehand that these can only be numeric values between zero and nine. Classifying other images like differentiating between golden retrievers and breaded fried chicken also falls into this category. Clustering, on the other hand, is all about grouping data together. These can be either continuous or discrete in nature. You already saw the example with the different species of plants. Clustering relies heavily on the number of factors that the programmer is interested in. Say you have an image taken at a high school reunion party. This is a high-resolution image with some hundred different faces. You want to cluster the data such that each face is its own group. Each group will be made up of distinct, plottable pixels. Given a mishmash of lyrics snippets, a machine learning system may cluster them into different genres and original artists.

Collaborative filtering is similar to regression, but instead of predicting future values, its goal is to fill in gaps in data. It strongly relates to research in recommender systems and algorithms. You can imagine some big content type services like YouTube and Netflix make extensive use of collaborative filtering to recommend things to users. Collaborative filtering takes data from collaborators or user agents who have similar tastes to yourself and attempts to generalize content they like with content that you might like. More generally, collaborative filtering is used to predict missing data in other systems like sensor and financial data.

Finally, reinforcement learning is used to teach a machine how to learn from the environment. Unlike the other types of machine learning, reinforcement learning doesn’t require large datasets to get started. Data still plays a pivotal role, but it is aggregated in real-time rather than over time. Driverless, for example, directly learn from their environments. If there is an accident, the system adjusts its parameters for the next time when a similar situation comes along. Arthur Lee Samuel’s checkers program would be considered a type of reinforcement learning. Recent advances have led to the creation of programs like Google’s AlphaGo Zero, which took the original AlphaGo program but generalized it to learn how to play Go without any dataset input whatsoever. AlphaGo Zero would go on to surpass the capabilities of its trained counterpart in just a matter of days.

As you can probably imagine, machine learning has several different applications in our modern lives. From facial recognition to stock market prediction and driverless cars, machine learning applies both principles of math and computer algorithms to simulate real intelligence. The common problems being solved by machine learning involve some type of regression, classification, clustering, collaborative filtering, and reinforcement learning. You could argue that there are a few more than these, but they are the most common ones. Anytime you hear machine learning in a headline or product description, it is probably about solving one or more of these problems.

Chapter 7: Neural Networks

If humans are intelligent because of their brains, and if brains work by creating neural connections called synapses, wouldn't it make sense to simulate these networks of connections in order to simulate intelligence in machines? Or at the very least, that is what early AI researchers thought. The sheer volume of connections in the human brain is what we owe our intelligence to. The average human brain has around one hundred billion neurons or ten to the eleventh power. These neurons can then connect to up to 7,000 other neurons – meaning the total number of connections is an order of millions of billions of connections. That is bananas. The beginnings of artificial neural networks (ANN) directly coincided with the study of real neural networks. In 1943, a neurophysiologist named Warren McCulloch teamed up with mathematician Walter Pitts to describe how neurons in the brain might work. They co-authored a paper in which they created a simple ANN out of electrical circuits. The ANN they designed used artificial or logical neurons called the McCulloch-Pitts neuron.

Inside the brain, a neuron works by receiving inputs, processing the information, and then transmitting it to other neurons. A neuron cell is made out of a nucleus that forms the cell body. From the cell body, structures called dendrites branch out like the arms of an octopus. Attached to the cell body is a long chain-like structure called the axon that is used to connect to other neurons. This point of connection is called the synapse and also looks like tendrils or branches used for connecting. The dendrite structure receives information and the cell body or soma processes it. The output then gets fired through the axon and into the synapse where the next neuron receives it. This is, of course, a simplified version of the true story, but it is enough for understanding artificial neurons and ANNs.

The McCulloch-Pitts neuron is, of course, purely logical. It doesn't make sense to talk about parts of a neuron as if they existed in real life. But it does make sense to talk about the parts according to their logical function. These artificial neurons are made up of two parts simply called *f* and *g*. First is *g*, it acts as the dendrite and receives some input, performs some processing, and passes it on to *f*. The processing can be a chain of Boolean operations that are said to be either excitatory or inhibitory decisions. A decision that is inhibitory has a greater effect on the neuron firing or not. For example, if the neuron is deciding whether to eat at a restaurant, an inhibitory decision would be something like "Am I hungry?" Obviously, if you aren't hungry, you won't make the trip. Less important decisions in the process may be "Do I crave fast food?", "Do I feel like going out?", "Does my car have enough gas?" and so on. These other excitatory inputs will not make the final decision on their own, but together they might. Next, *g* takes these inputs and aggregates them using a function. For the *f* to fire, the aggregate score of the inputs needs to surpass a certain value called the threshold parameter.

More specifically, an artificial neuron is a mathematical function. It has a number of inputs and an output. In the case of McCulloch-Pitts neurons, both inputs and outputs are expected to be Boolean values (true or false). It is also known as a linear threshold gate. The structure of the artificial neuron allows it to simulate logic gates. To simulate a logical AND operation, the neuron takes on a threshold parameter of three given three inputs. In other words, the neuron only fires if all three inputs are true. A logical OR operation takes three inputs, and the threshold parameter is one. Note that the logic gate can be a little more complex when adding inhibitory inputs. For example, a neuron with two inputs can form AND logic, but if one of those inputs are inhibitory, then the neuron won't fire. It will fire, however, if the inhibitory input is set to false.

The NOR and NOT logical gates can easily be derived from the previous examples.

If you are familiar with if-else logic in computer programming, you can probably tell that this scheme is essentially simulating long chains of if-else logic. However, the mathematical neuron model can “learn” the outcomes of decisions without needing to calculate each if-else statement. The logic is reduced to a simple function that outputs either true or false. This is also called linear decision boundaries. The artificial neuron splits inputs into two broad categories: positive or negative, fire or don’t fire. In an AND neuron with two inputs, this means that the only positive class happens when both inputs are true. Say that the neuron is deciding whether to go to bed. The first input can be “Is it past 11 pm yet?” and the second input is “Is tomorrow a workday?”. If both of these are true, then the neuron fire signifying it is time for bed. Conceptually speaking, the neuron has just learned what bedtime is – though 11 pm may be a little late for most people.

The McCulloch-Pitts neuron is an extremely simplified version of a neuron abstracted into logic. Other types of artificial neurons also exist. And just like neurons connect to others to form synapses, the same can be said for artificial neurons. That is, every ANN will use some version of a neuron as their most irreducible unit. Using the same McCulloch-Pitts neuron, we can imagine what a neural network may look like. Artificial neurons are organized into different layers that feed their outputs into other neurons. We already saw how one simple neuron AND gate can learn when it is time for bed. Imagine what kind of behavior hundreds or thousands of these neurons can achieve. Since McCulloch-Pitts neurons use Boolean logic only, the things that they can compute is a little simplified compared to what other artificial neurons can use. While they can only pass true or false values to the next neuron, others may pass weighted values. Though the principles remain the same, neurons only fire if a certain threshold value is passed. Since an ANN can have multiple inputs going into multiple neurons at the same time, these inputs are said to propagate or cascade down the network. Instead of returning a true or false value, more sophisticated artificial neurons may trigger program behavior like steering an autonomous vehicle a few degrees to the left in order to avoid a pothole.

The story of how ANNs derive weighted outputs from neuron connections is a little more complicated. Two methods at the heart of many ANNs are called backpropagation and gradient descent. The algorithms that make machine learning and neural networks viable come from the branch of applied mathematics called optimization. Mathematical optimization focuses on the selection of the best element from a list of alternatives and the criterion necessary for making that selection. When you train an ANN using supervised learning, you need to use something called a cost function that calculates the error rate between what the ANN predicted and what the correct answer is. The cost function is really the aggregation of individual loss functions, which calculate the error rate for single training examples. This relates to an algorithm called gradient descent which is used during the training phase of an ANN. The algorithm’s purpose is to find the values of parameters that reduce the cost function used by the ANN as much as possible. In other words, cost functions, gradient descent, and backpropagation are really the bread and butter of ANN machine learning. Without them, there is no indication that the model is learning from the data you supply it with.

You probably know what a gradient is if you have ever messed around with a graphics editing program. They are used to mesh two different colors together in varying degrees of intensity. A mathematical gradient sort of does the same thing, but they measure how much outputs change depending on slightly modifying inputs. A gradient can be calculated as a sort of slope. In machine learning, the higher the slope, the faster an ANN can learn. And if the gradient slope is zero, then the ANN doesn’t learn. The simplest analogy to understand mathematical gradients is a blind hiker going up a mountain or hill. His objective is to reach the summit with the lowest number of steps. The peak is relatively flat with a small slope, but the base of the mountain has a large slope. At first, the hiker can take longer strides up the mountainside to minimize the number of steps, but as he approaches the top, he takes smaller strides because he wants to arrive at the summit and go past it. In order words, it is easier to cover more ground when the slope is high. The amount of height you climb relative to the length of your strides is high, but it goes lower the higher you climb. If you may recall high school math, a slope of 0 indicates a horizontal line. A higher number indicates a steeper degree or angle of tilt.

A gradient descent then is going down the mountain towards a valley or the bottom of the function curve (if it was plotted). It is a minimization algorithm, so this makes intuitive sense to go down. If you have a machine learning problem with a cost function having two parameters W and B, gradient descent will try to find the values of those two parameters that result in the lowest value for the cost function. This means that the overall error rate of the neural network goes down. Another concept called the learning rate is a measure of how quickly a gradient descent should go down. A higher learning rate means that the descent may offshoot the local minimum by a lot, and a lower learning rate means that the descent will eventually reach the local minimum, but it will come at the cost of time and performance. Note that arriving at the best local minimum is synonymous with the system achieving the best accuracy. A higher learning rate, then, can lead to inaccurate results. A good approach is to try to find a rate somewhere in the middle between fast and slow.

Backpropagation is simply finding the error rate, loss function or cost function through gradient descent and applying it to the weights of artificial neurons in the network. In simpler terms, backpropagation is a mechanism that takes an error rate and modifies the program to learn from it. When a handwriting recognizing program classifies a weird looking zero as a nine, backpropagation adjusts the weights so that in the future, weird looking zeros are classified correctly as zeros. In practice, this is a little more complicated, but the general idea remains the same. A machine learning system can only learn if mistakes are corrected. Without something to propagate corrections into the neural net, there would be no learning. No matter how many data you supply or for how long you run your training sets, the system would never get over the zero or nine slumps without modifying neuron weights.

Just as there are different types of artificial neurons, there are different types of artificial neural networks. Each has their own applications and methods for handling inputs and returning their respective outputs. The basics of convolutional neural networks (CNN) will be covered in Chapter 9 as they are part of deep learning. Another type called recurrent neural networks (RNN) use a structure where inputs don’t go directly from neuron to output. Instead, inputs can bounce around neurons to form a “recurring” learning pattern. One type of RNN called long short-term memory (LSTM) networks try to simulate memories with each logical neuron holding on to some piece of information along with the given input.

Though genius as they are, neural networks probably don’t mimic exactly how intelligence comes around in humans. They have been called the greatest algorithms invented in our lifetime, but perhaps they are just that and nothing more. Certainly, most applications of neural networks belong to the “narrow” version of artificial intelligence rather than the general. This in part stems from the sheer complexity of the human brain. Even the most complex neural networks hardly approach the raw computing power of the human brain. And even if they did, they are likely trying to solve one or more of the five general problems of machine learning. They are not trying to simulate thought, nor are they formulating abstract ideas all on their own. It has also been said that just because the human mind is the most complex thing known to man, it doesn’t mean that man cannot create things that are even more complex. While this is a logical statement, such complexity has yet to be seen in modern machine learning techniques. These systems stem from mathematical complexity in a purely systemic view, but they do not compare to the biological complexity of the mind. If all it takes is one competent programmer and some open source machine learning package to start simulating intelligence, there is a marked absence of complexity there.

Neural networks certainly have their uses and may very well be some of the most important algorithms known to man, but they are, at the very core, simple mathematical abstractions. Recall that the most useful problems in machine learning are solved by supervision, having vast quantities of pre-tagged data the system can learn from. The human brain, in contrast, can learn to categorize things without any supervision what so ever. A young child even not knowing what a dog is can readily identify a dog even if they lack that knowledge. They can identify their parents without knowing what a parent is. Someone can argue that the human brain always uses a version of supervised learning because we get bombarded by sensory inputs daily, but this is all essentially untagged data. The equivalent would be to train a neural network on audio a child hears over the course of the day and ask it to identify the mother’s voice. All the neural network can do is categorize similar sounds, but it cannot “tell” what belongs to what. This is something the human brain does at an intuitive level.

Chapter 8: Reinforcement Learning

Traditional machine learning schemes of supervised and unsupervised learning are usually pretty static – meaning that they follow set principles of data aggregation, neural network design, and then training. No matter what type of machine learning techniques are used, the system ingests the data and learns from it. The data can change over time, but the system doesn't generate any data of its own. Reinforcement learning changes all of that. While still technically a type of machine learning, reinforcement learning goes a step further by adding a "software agent" that can learn from data derived from the learned environment as well as generating its own feedback. A software agent is simply a robot or autonomous program that is designed to mimic the properties of agency in human and animal actors. This software agent acts as the main source of intelligence in the program. Instead of being fed a correct set of outputs through classified data, the agent learns through the simulation of reward and punishment. Because of this, reinforcement learning is considered a third paradigm in the machine learning sphere following supervised learning and unsupervised learning. As such, there are no data sets for training the agent. Instead, they are said to learn from environments and their own feedback systems.

You can imagine one possible software agent being a computer program tasked with finding its way out of a maze. Confronted with the same problem, a computer science student might use a pathfinding algorithm to find the exit, but a software agent operates based off little fundamental principles. It doesn't quite know what a maze is, but it may be programmed to seek reward and avoid punishment. A possible reward in a maze, for example, is moving to a previously undiscovered cell and a possible punishment is going over the same cells. These reward-punishment systems allow the agent to eventually navigate its way to the exit of the maze – though the agent still doesn't understand moving from one cell to another unless it is explicitly programmed. One possible form that these software agents are designed is through finite state machines (FSM). Rewards and punishments are then based on individual states that the machine encounters. For a maze solving program, a possible negative state is getting stuck in a dead end. The program will learn that it is to avoid them in the future like a child remembers not to touch the stove burners.

To begin the process of changing states, the program often needs a randomness function or stochastic process that simulates decision making in an unknown environment. There are very little scenarios a human can encounter where they have a completely blank slate. If we encounter a maze, we intuitively think about ways to traverse it. If we get put in a new and frightening environment, we consolidate knowledge from previous situations that may help us out. Think about those escape room puzzles you can solve with a group of friends. Once you are in there, presumably locked out, you have to find your way out by solving the clues that you are given. But no matter what you do in that situation, you generally have a purpose for doing it. Very seldom do we say that humans do things randomly. This aspect of randomness is a central problem in reinforcement learning because intelligence is supposed to be modeled after purpose, not the roll of a dice. Unlike traditional machine learning, reinforcement learning is more akin to the study of decision-making. It borrows concepts from several disciplines including computer science, economics, neuropsychology, and mathematics.

You may have heard of "positive reinforcement" if you have ever taken a psychology class. Notice that positive and negative reinforcement sound similar to reinforcement learning. That's because they directly influence the field of reinforcement learning. Why does an animal or human do something – anything at all? Well, one possible motivator is the experience of pleasure or personal gain. On the flipside, avoiding negative outcomes strengthens behavior just as well, if not better. Reinforcement learning takes these concepts a bit further because it seeks to find optimal decisions to solutions. A software agent can visit each cell in a maze and still find the exit but doing so is tedious and inefficient. A better solution is to avoid dead ends (negative outcomes or punishment) so that the agent doesn't keep stepping in cells it has already visited.

Because of these reasons, reinforcement learning belongs to a subset in artificial intelligence research that seeks simpler, general principles. If the nature of intelligence is simply smart decision-making, then a sufficient breakthrough in reinforcement learning might lead to the creation of the first general intelligence systems that reach or surpass human intellect. Whether reinforcement learning is suitable for that end is debatable, but it clearly differs from the other types of machine learning. Reinforcement learning is limited because the power of neural networks and their algorithms are also limited.

Chapter 9: Deep Learning

If neural networks approach the way that human's think, deep learning takes the idea a step further. Neural networks and artificial neurons have a long history dating as far back as 1950. But when they were first introduced, computing power was limited, and ANN was looked at like research toys rather than hardcore business facing algorithms. When computing power improved, ANN received renewed interest from AI researchers as well as large internet companies. We are just now emerging out of the AI research winter that persisted into the 1980s and 90s. Part of this is due to computing power, and the other part is due to the introduction of the web and the massive amounts of data it generates. Today, deep learning is at the forefront of AI research and continues to make progress, helping the field thaw from the cold.

Two things limit an artificial neural network. First is the computing power necessary to simulate layers of artificial neurons, and the second is a combination of available data and feature selection. Even the most powerful ANN clusters today are still orders of magnitude behind the raw computing power of the brain, for example. The introduction of powerful graphics processing units (GPU) for machine learning purposes greatly increase available computing power across the board. GPUs are inherently faster than CPUs because they tend to prioritize smaller, more efficient cores compared to the CPU's powerful but bulky ones. They also allow for multi-threading of computational tasks and can easier perform floating point arithmetic (decimal numbers) than CPU's. Though GPU's were intended for rendering 3D graphics at hundreds of frames per second, they have been adopted by the AI community for processing large deep learning projects.

So what exactly is deep learning? As the name applies, it relates to creating additional layers of depth that traditional ANNs do. The argument goes that if the brain is made up of layers and layers of neurons, how is it that flimsy ANNs with singular layers are capable of simulating intelligence? These are sometimes called "shallow" neural networks to differentiate from those with multiple layers. Now that GPUs are extremely fast and getting better every year, deep learning doesn't require entire datacenters or neural net clusters to train models.

Returning to the human brain motif, deep learning takes after the tendency for complex ideas to fire deep in the folds of the brain, rather than at superficial levels. Recognizing edges of pictures and tiny details fires neurons closer to the surface of the brain, while recognizing larger constructs like a person's face fires deeper. More layers equal better, more intelligent systems. Information passes from the input neurons to additional hidden layers that also pass those inputs into each other. The more of these hidden layers, the better the results of machine learning. This is why deep learning is capable of tackling problems in artificial intelligence that shallow learning has traditionally lacked behind in. This includes computer vision, voice recognition, and language processing.

The true power from deep learning comes from its non-linear processing of features. Traditional machine learning techniques mostly use linear models and suffer from the feature engineering phase. With deep learning, features don't need to be picked out by a field expert. Instead, many different features are picked per model, contributed to the overall complexity of the neural net. A traditional classification of something may have used two or three features, but the deep learning equivalent is to use as many as the data affords. For example, to detect whether an object on the road should be considered a vehicle obstacle, a shallow machine learning system can use the shape of the object and its speed as factors. Such a system may perform well in the short run, successfully identifying different makes and models of cars whether in motion or parked. However, the system may encounter some unspecified behavior like a large carnival float moving relatively slowly around many pedestrians. Using shape and speed alone would not be enough to classify it. In contrast, a deep learning system may use several different factors in addition to shape and speed. As baseline inputs, shape and speed get passed to the deeper layers where they may also compare proximity to a road, the presence of pedestrians, orientation, distance from the camera, and so on. These additional factors will take longer to train the model and more computationally expensive, but it will be better at identifying vehicles in the long run.

Consequently, the scope of traditional machine learning has its limits. There is a point where introducing more labeled data doesn't result in better performance of the system. With deep learning, though, adding more data directly leads to better performance. It is mainly an issue of scale. One can scale well to a large number of inputs but the other cannot. It is no wonder that data obese companies like Facebook and Google use it. Their main value as a company comes from the data that they acquire. Deep learning allows them to exploit it, gain insights, and ultimately profit from it, and the reason why deep learning algorithms scale so well is that they are more conducive to analog like data that span many features. Data like images, audio recording, unlabeled text, and video footage are very different to work with than neat tabular data. These types of data are particularly good at forming hierarchical representations of features. Since the features don't need to be chosen beforehand, deep learning algorithms can learn to form classes of features on their own. Higher level learned features will be defined in terms of the lower level learned features. Returning to the vehicle identification example, a low-level feature may be some small defining aspect of the car like the rear windshield. A higher-level feature is a collection of these, like bumper size, indicator light positions, and license plate area, used to identify different makes. A car with a higher windshield and blocky appearance may be an SUV class, whereas something that is close to the ground may be a sedan class.

Deep learning neural networks work a bit differently from regular ANNs. One class of these networks are called convolutional neural nets (CNN) and are used primarily for image recognition. Like other neural networks, they are designed after biological processes in the brain. Animals use their visual cortex to perceive light through individual cortical neurons. Each of these neurons corresponds to receptive fields that overlap in the retina. CNNs work in a similar fashion. They consist of an input and output layer plus additional hidden layers in between. The hidden layers use something called convolution to process their inputs. Put simply – convolution is using two distinct functions to create a third function that expresses how the first one affects the second. Convolution is used to group pixels together from the beginning so that the net already has an idea of how the big picture fits together. It is easier to form a hierarchical structure of features as well. These networks first recognize small edges of the image as the smallest possible feature. Each layer progressively adds another edge or midsection to the hierarchical data representation until the whole image is learned.

Despite their increased accuracy, deep neural networks suffer from a number of drawbacks. Just because deep learning is state of the art, it doesn't mean it should be generalized to every conceivable machine learning problem. For many tasks, shallow neural networks are the preferred option. However, large companies like Facebook regularly employ deep learning because they have the requirement, data, and computational resources to perform it. Facebook recently said that it uses some billion digital images to teach its deep learning systems. Smaller players in the AI scene do not have the processing power to work on that scale. But then again, Facebook is one of the biggest companies out there. Google demoed how powerful some of its systems are a few years ago. Their system purportedly consisted of one billion connections. It was trained using YouTube data and could accurately recognize cats in the videos, yellow flowers, and other images. It is interesting to note that none of these features were selected or programmed outright. Their neural networks identified them through the hierarchical data representation. Furthermore, it could recognize between 22,000 different categories of images with some 17% accuracy. That level of accuracy is quite astounding once you consider the number of categories and how the system learned without any human first labeling the data. This accuracy could be increased to 50% if the number of categories was lowered to 1,000.

Today, if some artificial intelligence system is at the bleeding edge, it is probably using deep learning. Virtually all of the big Silicon Valley tech companies are using it. When you use Google Translate on some arbitrary string, you are using a deep learning system. Every time you fire up your Amazon Echo to speak with Alexa, you are using deep learning. Google uses it to tailor your search experience to fit your personal interests. Over time, it has developed a database of knowledge dubbed the “Knowledge Graph” that contains some 570 million different entities and 70 billion facts. It is used along with Google Search to more accurately represent the data that a user may be looking for through their queries. For example, if you look up the name of a past US president, Knowledge Graph accumulates the relevant data and displays it in the sidebar. These small snippets of relevant data are gathered from sources across the web like Wikipedia and the CIA Factbook. Google says the information provided through the Knowledge Graph is capable of answering one-third of its 100 billion monthly user queries. And if you are ever lucky enough to ride in a driverless car – yep, it’s also thanks to deep learning technology.

Chapter 10: Recommender Systems

Recommender systems are used by all the major tech companies, especially those that are content leaning. You can be sure that YouTube, Facebook, Netflix, and others actively employ them into their products. They are designed such that the more you use these services, the more they can recommend you things that you may want to watch. This increases the average time spent on their site or service because the consumer is shown appealing video after appealing video. It also saves the consumer from performing a search to find what they are looking for. Who among us today logs in to YouTube and never touches the search bar? Many people can limit their YouTube session to simply watching videos on their recommended section, or another video that pops up in the sections underneath. As long as you are logged into your account, these videos are tailored to meet your interests. Even if you aren't logged in, YouTube increasingly uses your IP address and other browser agent details to tailor the homepage based on your previous viewing habits. Recommender systems are actively being researched because they add value to the company without needing to change anything. All the content is already there – these systems act as a force multiplier on the profitability of that content.

For a company to “know” who their users are and what their interests may be, they use filtering algorithms that compare users along system-wide user profiling techniques. The most common of these algorithms is called collaborative filtering. Another type is called content-based filtering. They achieve similar goals but differ in how they are implemented. The differences can be gleaned from two online music streaming services, Last.fm and Pandora. The playlists or “stations” generated by Last.fm use collaborative filtering to find out what other users with similar music taste are manually adding to their libraries. A brand-new user to Last.fm will only have content that they search for or actively listen to. After a few days or even hours of listening to their favorite music, the user will receive randomly inserted music into their libraries. This is the result of the collaborative filtering technique looking up other users who listened to the same artists and finding out what other music they enjoyed as well. If it all goes according to plan, the new user will be satisfied with the recommended music and keep listening to it. On the other hand, Pandora uses a content-filtering approach. They gather song attributes from their patented Music Genome Project database to find other songs and artists that overlap in their attributes. The database stores a mind-boggling 400 different attributes per song. This hodgepodge of different but similar music can then be refined by the user making a playlist. They simply “dislike” a song and the algorithm will deemphasize the songs attributes from the song selection process. The more a user dislikes, the more accurately the system can recommend the music they want to listen to. Likewise, if the user “likes” a song, then the algorithm looks for music with the same attributes that were liked. In Pandora's view, all music can be boiled down to these 400 attributes.

With the current information glut online, it is no wonder that companies developed these systems. In theory, recommender systems are a win-win solution. The customer doesn't have to sift through endless amounts of data nor do they have to suffer from information overload. The service provider, in turn, gets more profit. The main area of contention with recommender systems is that they might leak user information. Efforts to anonymize data in the past have met their fair share of criticisms by security and privacy advocates. Netflix was actually sued by a small group of individuals who learned that their anonymized data used in the Netflix Prize competition could be cross-referenced with free online resources to reveal their identities. Netflix went on to settle for an undisclosed amount. Other concerns in this space center around discrimination and possible misuse of collaborative filtering data. If this data isn’t anonymized, it can be linked to individual user accounts, meaning the company has private knowledge of the things you like. This, in turn, can fuel shady targeted marketing schemes.

The future of these systems points in a clear direction: every increasing level of personalization. Currently, recommender systems are isolated into the content world. Users like content, they can be recommended content and thus will want to use the service more. However, what if these systems could be generalized to areas that don’t focus on content? This technology combined with the internet of things, for example, could create the ultimate personalization climate for consumers. Recommender systems could connect to the user’s smart fridge and tap into a vast network of food preferences across all smart fridge owners. The smart fridge app could then recommend the user popular food items based on several attributes. Some of these apps could automatically place an order for them as well. Pandora uses 400 of these attributed to songs; a similar system could be developed for food. You can probably already imagine some of these attributes. Carbohydrate content, sugar content, keto friendly, diabetic food, and several others could contribute to healthy customized diets.

Chapter 11: Robotics

Robotics is an interdisciplinary field that combines elements of engineering, science, physics, computer science, and electronics, just to name a few. Robotics concerns the design, implementation, programming, and maintenance of robots. A robot can be any physical system that is designed to perform an action or series of actions with varying degrees of autonomy. Robotics has a long history and has probably been in the human psyche ever since the first Greek stories of automatons were being told – though today's robots are a little different from the ones you learned about in Chapter 2. Modern technology allows robots to be smarter, more lifelike, friendlier, and overall more useful than ever before. Many believe that humanity is headed for a new industrial revolution that will be spearheaded by the mass adoption of these systems in virtually every economic sector. Besides the obvious sort of industrial robots that you find in car manufacturing plants and refineries, there are also robotic systems being developed for retail, hospitality, transportation, telecommunication, medicine, and food services.

Robotics and artificial intelligence often get lumped together, but they are both distinct things. While some aspects of robotics lend themselves to use artificial intelligence techniques, the field as a whole is not preoccupied with simulating intelligence. Robotics can also be divided into several subfields that specialize in certain applications of robotic principles. On the industrial side of things, robots are used to handle tasks that are easy to delegate to a machine. Some of these tasks require high precision or repetitive motions that would be too difficult for a human to sustain for long periods of time. Others are designed for high throughputs like food packaging and labeling. These systems vary in how many behaviors they are capable of performing, their physical properties, and how they are programmed. Industrial robots come in different varieties, but one of the most common configurations is the "arm" type. These have similar designs but perform different tasks. They also have different degrees of freedom that allow them to move differently. Kinematics is very important for these types of robots because the mathematical equations determine how the joints on the arms move.

Interest in humanoid robots severely lacks with industrial robots regarding market penetration and usefulness. While research in androids or human-like robotic systems remains high, the role of the humanoid robot remains a techno-curiosity more than a commercially viable business – though one day it is conceivable that they will be bought and sold as housekeepers like the robots in *The Jetsons*. There already exist housekeeping aid robots like the Roomba that has been available since 2002. The Roomba is manufactured by a robotics company fittingly named iRobot. Rodney Brooks, a professor in robotics at MIT, co-founded the company in 1990 with fellow classmates. Since then, Brooks has gone on to start another company called Rethink Robotics in 2009, most known for its creation of the Sawyer and Baxter collaborative robots. Baxter is fundamentally an industrial robot with industrial capabilities but with a bit of imagination. Unlike most industrial robots that usually resemble a mechanical arm, Baxter has two main limbs and an animated LCD face. It was designed to perform mundane tasks on an assembly line.

Baxter is a humanitarian take on the industrial robot. Instead of performing tasks rapidly and mechanically, Baxter uses sensors to become "aware" of its surroundings. The LCD screen will display a face according to the robot's state. It can also respond to changes in its environment like ceasing operation if it drops a tool or piece and is unable to recover it. One of the most intriguing facets of the robot is that workers can programme it in the facility. Where other industrial robots require an engineer to be configured through control systems, Baxter is programmable by “hand”. An unskilled worker can move the robot’s hands to perform a certain task, and the computer will try its best to reproduce the movements. This makes Baxter accessible to everyone on the production line, not just educated technicians. Baxter has been lauded as being safer than traditional industrial systems that do not pay attention to other factors outside of their immediate programming. Despite this, the company went out of business in October 2018 due to low sales. Many old guard manufacturing companies saw these systems as experimental and the technology probably not there yet. The robots Sawyer and Baxter are still used in research today. Some universities use them to teach students in robotics courses. More recently, researchers hooked up electrodes between Baxter and a human operator to directly transmit brain signals. Sometime in the future, the interaction between human and robot may resemble something that is as intuitive as simply thinking.

Most industrial robots do not require to sense things or be particularly smart. They perform a set task and are usually left alone. In contrast, humanoid robot research focuses on the ability of the robot to perceive the world around them and to interact with it. To be able to move a robot requires some propulsion system, like a series of actuators or electric motors. Another option is to use hydraulics or pneumatic systems. All robots need an energy source, either a battery or a direct connection to the current through a wall socket. To be able to see, a robot can be equipped with cameras, LIDAR, and various sensors. Computer vision is coupled with the cameras, and the sensors directly send environmental data like speed, position, balance, and so forth to the robot’s control system. If machine learning is just a combination of statistics and programming, robots are a combination of a great many other things. In both cases, neither approach a general intelligence like the *Star Wars* droid 3-CPO. If that is the case, what is the current state of the art in robotics? Robotics research is split into many areas of focus, with some people working on sensors, robot dexterity, human-robot interaction, autonomous motion, and so on.

Humanoid robots are projected to disrupt areas of retail, hospitality, and food service. The chances are that you have already seen the automated menu systems for McDonald’s and other fast food restaurants online. These systems are definitely being worked on, but may not see the implementation for various reasons. Some startups are even working on burger making systems that can perform virtually all of the duties of the average burger flipper at a fraction of the cost. The role of the humanoid robot in customer service is more obvious in places like Japan where robotics has entered the mainstream. Of all the industrialized nations today, Japan has the highest robot density by a large margin. Most of these robots are used in the automotive industry, but a number of them take on customer service roles. They can be found in select department stores greeting customers and in airports acting as luggage carriers. With an ever-contracting population and less unskilled workers willing to do these jobs, it makes sense that Japan has a high adoption rate. For many Japanese citizens, service robots serve to increase the quality of life and are a normal part of their daily lives.

However, there is still a long way to go for the average consumer in other countries to warm up to these robotic and automated systems. Some would even say that the need simply isn’t there. In 2017, Japanese deaths outnumbered births by 1,000 to one. The population dropped by a massive 264,000 people in one year alone. Other industrialized nations with rising unemployment rates and a stable population perhaps don’t need as many robots. There are also other concerns besides just population and unemployment for why service robots are yet to be seen abroad. Overcoming the uncanny valley effect is no easy feat, especially with older generations. In Japan, people are used to them, so the effect holds less weight – though the same cannot be said for older generations in other countries. Young people today are more comfortable using self-checkout lines and automated systems, but the older generation overwhelmingly subscribes to the importance of face-to-face interaction with their community. A big argument for service robot adoption in Japan is their large aging population. End of life care is synonymous with assisted living, loss of personal autonomy, and embarrassing mishaps. These are areas that service robots can directly address. For one, an older patient is less likely to feel embarrassed if a robot is cleaning after their bathroom accidents. If there is a strong human-robotic interaction, the patient may feel that the robot is a part of themselves like we think our phones are. The result is a net gain in personal autonomy, all things considered.

With a shortage of geriatrics and personal care professionals plus a global aging population not just in Japan, service robots are projected to rise. Less and less people are signing up for these important ends of life careers. The pay for these jobs is low when compared to other medical professions that require a similar amount of training. If nobody else wants to take care of the elderly, who will? Not every family is willing or able to care for aging parents. Additionally, the risk of malpractice and abuse in assisted living institutions is high. Taking care of the elderly is not an easy job by any means. Combine this with a meager wage, and you have the basis for unfair treatment. Every aging person deserves to be treated with dignity and respect. Whether service robots can render this kind of treatment is up for debate. And it is a debate that will doubtlessly unfold within the current century.

Would older generations be keen on using a service robot for their daily needs? To some, a robot may seem intimidating or impersonal. After all, robots are not intelligent the way that we are. They may be able to handle basic conversational skills and respond to our commands, but that is a far cry from genuine intelligence. There is also the question of whether these systems should have a human-like appearance or to remain obviously robot looking. Humanoid robots are more likely to create an uncanny valley effect, but it is not exclusive to human-like design. Any robot that acts sufficiently “alive” can elicit the effect. Research into the area of the ameliorating, the uncanny valley falls under human-robot interaction. We know from this research that certain robotic behaviors elicit certain human emotions like fear and uncertainty. For example, when a robotic agent gets too close or invades somebody’s personal space, there is a fear response. If a robot is just lounging around with no purpose, the human perceives it as daunting and even useless. Interacting with these systems can result in unanticipated behavior in the human. Humans are likely to ascribe personality traits to a robot even when these aren’t explicitly programmed. This factor of unpredictableness has lead researchers and designers to add emotive cues to their robots.

Where service robots suffer from the uncanny valley effect, utility robots do not. Virtually every major military across the globe is investing in robots to aid in the battlefield and in operations. A US-based company, Boston Dynamics, is working on robotic systems known for their mobility. The company made headlines with its design of a quadruped robot named BigDog for the Defense Advanced Research Projects Agency (DARPA). The robot has obvious military applications as a pack mule carrying ammunition and supplies for soldiers on patrol. The project was eventually scrapped because the system was deemed too loud to be used in live combat operations. These sorts of systems are mostly considered prototypes though. Other robots used by the US military include the Foster-Miller TALON family of remotely operated vehicles. They resemble little tanks or planet rovers with fully tracked movement. They are capable of using small arms fire to heavy machine guns completely remotely. Though TALON saw some deployment in Iraq and Afghanistan, its current use remains limited. Other fully weaponized systems are readily used like the General Atomics MQ-1 Predator drone and the MQ-9 Reaper. These belong to a class of aircraft called UAVs or uncrewed aerial vehicles. Predator and Reaper drones can either be controlled by an operator on the ground, or they can fly autonomously with the direction of onboard computers. These systems have faced increased scrutiny with high profile killings involving civilians and questionable rules of engagement. In 2011, under the Obama Administration, a 16-year-old American citizen of Yemeni descent was killed in a drone strike while eating at a restaurant in Yemen. It is unclear why the boy was targeted, but his father was a suspected al-Qaeda leader and was later killed in a similar strike.

The legitimacy of so-called “killer-robots” comes into question now and then when a new system is announced. At the heart of the debate is whether a nation should possess the kind of power to kill remotely without trial and in the case of 16-year-old in a country that is not at war. Commentators are quick to point out that a machine with killing power is capable of being hacked. They appeal to the “Skynet” scenario where autonomous weapons gain sentience and begin to exterminate humans. Such scenarios are mostly far-fetched science fiction and are not helpful in the debate. While we know that some of these systems are vulnerable to getting hacked, we have yet to see a hacked autonomous system cause loss of life and limb. If there ever is a case where one does, it may cause enough trouble that governments are forced to ban them. Still, others question whether robots going to war is even a good idea. If state-of-the-art machine learning algorithms can barely drive autonomous cares with a certain degree of safety, how can we expect a robot to accurate wage war? War has several more dimensions of complexity than driving down the street does. Deploying actual robo-soldiers seems like an unlikely use of military spending with current technology – though you can rest assured that major governments are researching them.

A “killer robot” is a blanket term that is also unhelpful. Should a Predator drone be considered a robot any more than a tank with an autonomous turret is? Sentry guns have been in development for several years now and are regularly deployed into the battlefield. The most common of these are called close-in weapon systems (CIWS) and are used to protect battleships from missile and aircraft attacks. A CIWS is basically a large caliber cannon with a high rate of fire that is capable of shooting down missiles at 4,000 rounds per minute or higher. Unlike your grandfather’s anti-aircraft systems that may have required manual sighting, CIWS use radar guiding on a rotating mechanical platform to lock on targets autonomously. Not only that, but once “live”, the weapon system can engage without manual input after locking on. Any missile or aircraft that is detected by the radar system is automatically fired upon. In the case of the Phalanx CIWS used by the US Navy, a single 20mm Vulcan gatling cannon is used with computer-like accuracy. The problem of targeting is reduced to a simple algorithm. It doesn’t require machine learning or any other fancy artificial intelligence techniques. If the radar picks up an object that is approaching the ship, it needs first to make sure that its course is directed towards the ship. If the approach of the object is directly heading towards the ship, then the system must decide to fire. If the object is in between the minimum and maximum velocity range, then the cannon fires. But if the object is too slow or too fast, the system does nothing. Both the minimum and maximum velocity range can be programmed by the operator.

CWIS, like the Phalanx, are considered autonomous weapons, yet they use very simple tech that has been around since the Gulf War. However, by today’s standards, this is still a “dumb” cannon. The imperfect nature of the weapon has been demonstrated in a few disastrous training exercises. In one instance, the system successfully shot down a dummy drone target, but fuel and debris from the explosion damaged the ship and crew members. In another instance, the dummy drone was shot down, but the system reengaged it as it was in free fall, inadvertently sending rounds in the direction of an opposite ship and injuring crew members. This behavior should be expected from its simple rules of operation. A more sophisticated weapon system could probably have avoided these incidents. But as a last resort missile defense system, the Phalanx does a relatively good job against isolated missile strikes. Other autonomous, radar-guided systems are common throughout the world. Again, these don't need complicated machine learning schemes to identify targets and engage them. Israel's iron-dome system of missile defense is hailed by some commentators as the most advanced missile defense system in the world. These systems demonstrate some degree of intelligence, but really all they are doing is calculating velocities.

All things considered, the future of robotics is bright. Demand for robot systems will likely rise with uneven proportion to robotics professionals, creating well-paying jobs for those who are interested. There will be a glut of these jobs and a shortage of robotics designers and engineers. At the same time, lower-skilled jobs like robot technician and repairmen may emerge. Robotics will infiltrate other areas of tech like the cloud computing space and the internet of things. Robots connected to the cloud will lead to the creation of a cloud robotics marketplace were bidders can remotely program robots to perform certain tasks. Just like Alexa skills, these specific robot programs can be bought and sold in an open marketplace. Not only this, but robotics is taking on increasingly more important roles in business structure. A new executive position called chief robotics officer (CRO) will emerge for bleeding-edge organizations that make heavy use of robotic systems.

There are split views about just how far the field or robotics can go when it comes to automating human tasks. Whether or not the introduction of these technologies leads to technical unemployment and whether they should be regulated is up for debate.

Chapter 12: The Internet of Things

Imagine a sea of connected devices that can all communicate with each other and transmit information. The internet of things is one of those buzzwords that gets thrown a lot, but that nevertheless has a relevant name. The internet is short for “network of intra-networks” or simply network of networks. A company or university has their own network of connected computers that may or may not be connected to the internet. This is called an intranet. When you connect those intranets through a common routing protocol as the internet does, you have digital communication on a massive scale. The internet of things then is a network of connected devices. Smaller and smaller network adapter cards allow devices to be smaller and fit virtually anywhere. To give an idea of how small these devices can get, a Bluetooth low energy weather beacon is about 30mm in diameter, 10mm thick, and weighs about 7 grams. Besides being minuscule, this device has a low energy usage and depending on configuration, can last for a couple of years before needing a battery replacement. Devices such as these have a small transmission range depending on what wireless technology they use to communicate. But since they are small, inexpensive, and massively produced, many devices can be laid down in an area of interest to form a relay network that can effectively increase the range of transmission.

The internet of connected devices is set to explode in the coming decades both as these devices become cheaper and as solar energy cells improve. The Internet Protocol (IP) under IPv4 namespace for IP addresses allow for a maximum of 4,294,967,269 addresses. IPv4 uses 32-bit addresses about 232 bits in total size. There are only so much different combinations of addresses that the protocol affords. However, the much newer and improved IPv6 protocol uses 128-bit addresses and as you can imagine offers orders of magnitude more addresses. In fact, there are about 10 to the 22nd power addresses, which is a mind-boggling number. Part of the reasoning behind the implementation of IPv6 was that the world was running out of IP addresses to use for websites. Another reason was that the internet of things adds significantly more connected devices that all require unique IP addresses. With IPv6, the world is ready to stomach the burden of the internet of things.

At a consumer level, the internet of things penetration is high in recent years. The so-called “smart home” revolution has seen an introduction of smart devices, thermostats, garage openers, toasters, microwaves, televisions, and refrigerators that are all connected to the internet. These devices usually come equipped with a common interface or dashboard that is configurable through a smartphone. Why would anyone want to buy a common household appliance that is connected to the internet? It’s a good question, one that marketing teams all over the world had to grapple with when these products were being designed. The benefits of an internet connection are obvious for some appliances and less obvious for others. A smart fridge that comes equipped with a barcode scanner can keep an inventory of the household food supply, as well as note when products are about to expire. The complementary smartphone app can then maintain a database of all food purchases over the course of the year and offer basic analytics. More technically, inclined users may seek a way to export their data for further analysis. As you can imagine, these products tend to lean on the expensive side and are marketing towards middle and upper-class households.

A more general definition called household automation refers to the devices, technologies, and control systems that aid in the home economy. At the most basic level, you have things like garage openers and clap-on lights. Somewhere in the middle, you have lighting systems, home theatre systems, and temperature control. At the high end, you get things that are only limited by the DIY spirit of the owner. A capable enough individual, say an engineer, can conceive of an automated pet feeder system that only requires food to be replenished every once a while rather than at every meal. Consumer irrigation systems can also come with a “smart” component, possibly a dashboard interface with options for setting up the watering frequency and so on. FarmBot is an automatic home gardening system that requires virtually zero effort to maintain. It mimics the architecture of a CNC milling machine but is instead equipped with trowels and seed dibblers. While the basic kit costs upwards of $3,500, the company maintains that the system is 100% open-source. This falls in line with the DIY ethic behind many the internet of things products. Sales of home security systems have also been rising in recent years. Many technology companies are taking full advantage of the internet of things revolution in security spaces and consumers have multiple brands to choose from. Smart assistants like Amazon Echo and Alexa have also gained popularity. They aid in the home automation process by listening to user commands and executing some function. This may be to surf the web, create a shopping list, play music, and a myriad of other things. The user can link up “Alexa skills” that are the programmable scripts that the devices run after each command. These skills are readily searchable online, and anyone with the programming knows how to publish them.

Home automation is especially relevant to the aging population as the needs of the elderly are myriad when living at home. For some, systems may delay the need to be admitted to a healthcare facility. However, with current technology there is only so much these systems can achieve. Stairlifts for those who use wheelchairs have been around for years. Home automation then isn’t a new thing, but it certainly has gained renewed interest with the internet of things technology. An application of the internet of things principles to home automation for the elderly has many benefits. One of the most important features that many elderly homeowners need is an alert system that can call emergency services if they are incapacitated. Non-emergency alerts like reminders to take medication and set up doctor’s appointments are also helpful. The true use of the internet of things methodologies would include a smart device that is worn on the wrist or chest that sends information to the patient’s doctor about heart rate and blood pressure. Recent advances in smart fabric technology allow this functionality to be embedded into the very clothes that the patients are wearing. In the future, these systems will be able to integrate with robotics for additional assistance. Domestic robots will prepare meals, clean after the patient, and assist with doing chores.

Another novel family of the internet of things applications belongs to the industrial sector. Smart sensors can relay information about weather conditions, equipment status, and logistics. RFID technology can be used to track stock keeping units (SKU) that move from warehouse to warehouse. The so-called “industrial internet” combines networked devices, big data analytics, and real-time updates. These always-on metrics are relatively cheap for large companies like General Electric to implement and yet they provide a wealth of information. Areas where the industrial or manufacturing process cause inefficiencies can easily be spotted by their operators. Once a bottleneck is identified, the necessary changes to correct them are given center stage. Increasing the number of connected devices lowers sunk costs in the form of productivity losses by optimizing the entire process. When you have all the data in the world, you have options. Though General Electric remains an industrial internet powerhouse, the concept has seen slow adoption globally.

The internet of things devices have also been proposed for the creation of “smart grids” for common utilities. The idea is to use monitoring devices to gauge the level of electricity needs so that the grid can move resources to areas that need it most, meanwhile boosting efficiency. This enables two-way communication between the utility provider and the consumer. Smart grids ultimately lead to lower electricity rates for the consumer and general availability for all. A hidden benefit of transforming power grids into smart ones is that old equipment is replaced with the new. It is no secret that US infrastructures are slowly crumbling away. Installing new networked devices gives policymakers the excuse for finally getting rid of troublesome parts. A newer grid is a safer, more reliable grid. A smart grid is an informationally rich way to do utilities. The customer will have access to smart meters that can be accessed through their smartphones whenever they wish. There they can see their monthly bill, usage rates, and opportunities to save. A smart grid prioritizes resources for high-demand peaks, but at a greater cost. This allows thrifty customers to save their most demanding electricity usage for low-demand periods, saving them money in the process. Another benefit to smart grids is higher availability for vehicle charging stations. As it stands, the US power grid is not ready for the switch to a mostly electric car society. Instead, electric drivetrains are only practical in certain high-density areas where charging stations are readily available.

Since the internet of things technologies generates massive amounts of data, existing machine learning systems will only get smarter. If a company is already using sensors and other connected devices to harvest data, you can be sure that they are using a data lake or data warehouse solution to store it. Real-time analytics requires high bandwidth data ingestion engines to parse the information as it is being relayed to servers. Big data is only projected to grow alongside with the adoption of these networked devices. The internet as it stands right now generates a massive amount of data. Social media apps like Snapchat and Instagram upload thousands of user-generated photos and videos every minute. Facebook, LinkedIn, Quora and others generate data through user posts, as well as in-house and third-party analytics for each user. Virtually any high-traffic website is using third-party tracking and data harvesting add-ons. These record things like mouse movements, keystrokes, and clicks on advertising banners. Basically, any action that a human agent does on a logged in account can generate data for various companies.

If you add to this the internet of things, which generate different types of data formats depending on the wireless technologies, the global data glut shoots up. For example, the World Wide Web uses various data formats that are well known to data scientists. These include JSON, CSV (comma separated values), and XML. Many devices that are connected through web protocols like HTTP communicate with these types of columnar data. The number of connected devices is expected to reach 31 billion in 2020. That is almost 30 times the number of people who are currently online. The amount of data these connected denizens generate was somewhere around 2.5 quintillion bytes a day in 2017. We can expect the field of machine learning to grow, with new techniques and algorithms added to the already diverse repertoire of narrow artificial intelligence. Because connected devices can be virtually anything, it is difficult to say in what ways they will be implemented. We do, however, have a general sense in where things are going. The most important principles of the internet of things technologies are sensing, communication, and relaying. Therefore, the internet of things will be used to create large-scale communications systems of small devices.

A smart city is a type of urban metropolitan area that uses the internet of things connectivity for analytics, data optimized infrastructure and utility delivery, and transportation networks. A data orientated city is one that utilizes smart grid technologies to boosts energy efficacy, has an open attitude towards civic participation, and has optimized fleets of public transportation. Citizens can download an application on their phone that gives them their own personal dashboard into the city in real time. They can read metrics, look up transportation schedules, traffic alerts, and several other functions. Technologies will allow them to find empty parking spots, report potholes, and gauge the human density of their favorite hangout spots. Most of this functionality will stem from connected sensor networks. Vehicular networks, including in-vehicle and vehicle to vehicle communication, will make driving easier, safer, and will pave the way for fleets of autonomous driving vehicles. A modern vehicle of the year is equipped with countless ECU (electronic control units) that each communicates with each other in an in-vehicle network. These are used for on-demand vehicle diagnostics, conformance to emissions control standards, and even breaking systems. Vehicles are increasingly being computerized to the point where models that use drive-by-wire are common. These models use a type of electronic steering that obviates the need to have a mechanical steering column. Vehicle to vehicle networks allows for efficient throughputs in busy roads. They can even be configured to detect accidents and avoid them in real time. If cars can communicate their positions, speed, and direction, then deterring an accident is a simple means of the car braking or steering by itself when the driver isn’t paying attention.

Just like the other applications of these technologies discussed in previous chapters, the role of regulation and legal frameworks cannot be stressed enough. Even if the technology is there, it doesn’t mean that there exists a solid policy for implementing them. Driverless cars are a regulatory nightmare, so are vehicular networks that alter the behavior of crewed vehicles. Something that is brought up time and time again with connected devices is the availability of secure communication. For many in the security industry, connecting vital infrastructure and private networks poses a security risk that outweighs the potential benefits. Since this is a relatively new space with different wireless technologies, potential attack vectors are everywhere. It is difficult for the internet of things to maintain a strong security posture because there is greater overlap between physical and wireless intrusion. Besides wirelessly attacking a sensor or vehicular network, a potential attacker can also go to the physical site where the sensor is located. In industry terms, these are called physical layer and networking layer attacks respectively. Since these networks use many individual nodes, the attack surface is only increased. The sheer number of consumers the internet of things devices connect to the internet poses the threat of virus proliferation. The threat of massive scale denial of service attacks is a commonly cited fear by security researchers. Imagine a zombie network of millions of domestic appliances like toasters, refrigerators, speakers, and thermostats that can use web protocols to inundate legitimate web services with communication requests.

Only time will tell if the internet of things security posture can be hardened to the point that regulators are more willing to accept them. Continued research in the safety of vehicular networks will have to improve for municipalities to give the go-ahead on roadside vehicular sensors. The same can be said for smart grids. All across the globe, experiments in smart cities are being conducted as you read this. A simple Google search can point you in the right direction towards such a city near you. Whatever the case, you can expect to hear more about the internet of things’ security, implementation, vehicular networks and regulation in the coming decades. Any new technology is usually slow to be adopted, but when it is, it has the potential to revolutionize society.

Chapter 13: Why AI is the New Business Degree

Once upon a time, the most popular college degree was a business degree. If a student didn't know what field of study they wanted to go in but still wanted a decent job, the business degree was the way to go. Now, there is an overabundance of students going for their MBAs but not enough going after degrees in artificial intelligence. While the business world is facing a saturation of business knowledge, artificial intelligence is facing a shortage of AI know how. There has never been a time in history where getting a computer science degree focused on AI methods has been more lucrative. Many top tier schools are now offering new degrees that focus on AI and the gamut of robotics rather than the general computing knowledge taught in traditional computer science programs. This means that there is market demand, as well as plenty of candidates interested in the degrees.

It is no wonder why AI is quickly changing the face of business in multiple sectors. Technology related jobs are increasingly asking that their candidates have a solid grasp of machine learning methods along with their other expected duties. Virtually every major AI player has already open sourced some of its machine learning libraries so that everyone from startups to large corporations has access to create AI programs with little upfront costs. Google released TensorFlow in 2015, and it has since become one of the most popular repos on GitHub, the definite open source authority today. Another open source library called PyTorch is used extensively at Facebook and Uber. While smaller players in the AI space benefit from hiring PhDs in the field, they already have many of the core infrastructures for creating neural networks provided by these free tools. The availability of these tools is further coupled with easy access to computational resources provided by cloud providers like Amazon Web Services and Microsoft Azure. A company no longer has to invest in a high-cost machine learning cluster of GPUs when they can simply rent as much processing power they need from the cloud. While these services are more expensive in the long run, they still make it easier for a company than building machine learning infrastructure outright.

AI tools allow a business to engage in machine learning that before may not have found a use for it. A relatively new trend is to use automated chatbot software on their customer-facing web pages for quick and easy information retrieval. A customer can ask for rates, available products, and other information by simply typing a few text commands. While chatbots are a bit behind in terms of realizing their full potential, many consumers are becoming more familiar with them. Some companies even hire human agents to fill the chatbot role until the technology exists to allow chatbots to perform these duties on their own. Another common application of AI technology is to predict customer behavior. The focus on analytics is at an all-time high for all types of companies, not just retail. Coupled with social media marketing and e-commerce, analytics drives better customer decisions for the company over the long run. There has also been a rise in purely "AI" focused companies that market a single product or line of products that have AI functionality. One of these is called Grammarly, an online service that uses AI methods to streamline the writing process. It provides editing for simple mistakes, writing pitfalls, and suggesting things the user can say to sound more professional. It is essentially a writing tutor that the user can take with them wherever they go. Yet another company called Stick-Fix uses AI to recommend clothing options to its customers based on a series of preferences. A user designates their price range, enters their measurements, and chooses a style they are going for, and the service sends them a box of outfits to try on.

Large companies may use AI to automate their systems. This is especially true in manufacturing and human labor-intensive jobs – though many of these jobs will take years before being fully automated. The recycling plant example given earlier in this point remains to be solved on a massive scale. Most automation systems first begin by aiding line workers rather than replacing them outright. While low-skilled workers are at high risk of losing their jobs to automation, the shift will not occur overnight. The most likely scenario is that when robotics is first introduced into new industries, they will work alongside human employees. This is already true in the car manufacturing business. In 2018, Tesla Motors was highly scrutinized by business leaders in their attempts to automate large portions of their Model 3 production. The result was an overestimation of automation capabilities and an underestimation of human worker capabilities. The company undershot how many Model 3s they could output per month using their heavily automated plant. This move was criticized by industry veterans, some of them calling it a "rookie mistake". The state of modern robotics is still behind the power of the mind and especially human dexterity. Mimicking the same micro-muscular movements used for manipulating tools in the human hand is a non-trivial problem to solve with robots. It will likely take decades before a machine matches the human level dexterity of a line worker with years of experience in their craft.

However, it isn't just line workers who are in danger of losing their jobs. Advances in AI, especially in speech recognition and language processing, threatens to displace the huge call center industry with automated systems – though that is also likely to be decades in development. Retail is seeing a transformation with automated systems on top of the already high online sales numbers. Fewer people are going out to shop. At least in places like Japan, retail facing robots are actively being invested in. Even those who enjoy a comfy white-collar job have reasons to up their education and technical skills. Automated systems for payroll, accounting, and balancing the books are under active development. In the future, even programmers will not be safe from the AI deluge. Systems are being trained to perform the most mundane programming tasks that are often handed to lower-skilled coders. This includes software testing and looking for bugs – though, with all things under threat of automation, these systems will likely be deployed to work side by side with the programmer rather than replacing them altogether.

Driverless technology is also on the rise. Even as you read this, you can be sure that several companies across the world currently have autonomous driving systems on the road somewhere, gathering ever more data to strengthen their algorithms. What will likely happen with driverless technology is that the ability will come first and policy second. Just because driverless cars are proven safe, it doesn't mean that they will automatically be introduced. There are far too many gray areas to see the mass adoption of these technologies soon. In 2018, the first-ever driverless car fatality was recorded in Tempe, Arizona, by a car owned by Uber. You can be sure that more such fatalities will follow until the technology is perfected and here lies the difficulty in policymaking. There has never been a time in history when machine ethics have been used to determine laws. Even if driverless cars kill several hundred people a year (and they probably will), lawmakers and insurance companies have to decide if it is preferable to killing several thousand. There is an increasing need for stakeholders to have this ethics conversation, as well as for common citizens to be informed regarding policy and current advances. The major deterrent to driverless vehicles hitting the mainstream will be the law, not the capabilities of technology.

The industry is currently hungry for people skilled in the top AI libraries. The highest paying positions are looking for PhDs and Master graduates, but a good portion of them are looking for anyone at all who are competent programmers. The need spans other industries besides pure software engineer as well. Mechanical and electrical engineers with knowledge of machine learning techniques will be in high demand for the coming years. Add to this a cursory knowledge in the internet of things, and you have a highly desirable candidate.

Chapter 14: AI FAQ

Question: Does machine learning have limits?

Answer:

Machine learning certainly has its limits. Even when advanced techniques like deep learning are used, these systems are limited by their data and feature selection. Deep learning is a little ahead of that curve because features can be selected automatically. However, machine learning can only be applied to the five general problems discussed in Chapter 6. If a problem lies outside of that problem space, then current machine learning will never be adequate for answering it. Simply throwing more data will be useless if machine learning doesn't align with the problem. Fortunately, or unfortunately, we haven't had machine learning robot legislators anytime soon.

Question: How will AI be used in the military?

Answer:

The US Pentagon has increasingly shown interest in the term "algorithmic warfare". That is the application of machine learning methods to the battlefield. It will most likely be used for target selection and prediction of enemy movements. Autonomous systems with "kill decision" capabilities will no doubt be developed from these general applications – though as previously mentioned, there is a huge backlash against the militarization of AI by the research community. These weapons may or may not gain unconventional status by the international community somewhere down the line. Even if they do, major powers will still pursue their creation.

Another possible application of AI will be towards cyberwarfare, the infiltration of nations through subterfuge, and the targeting of autonomous weapon systems of enemy states. AI algorithms will only add potency to cyberattacks by state actors, with AI systems possible taking over enemy drones and uncrewed gun platforms.

Finally, there is the looming threat of facial recognition used to identify targets. It would not be out of the ordinary to suppose drone attacks in the future will be targeted at domestic actors. The drone strike in Yemen first sent shockwaves through the international community because it affirmed its right to strike at targets over foreign soil. We have yet to see such a strike inside the country.

Question: Should I be worried about losing my job?

Answer:

The short answer is no. Unless you belong to a certain market vertical where automation has already phased out many jobs, you are probably safe for the next ten years or so. If you are in a profession that is actively an area of AI research like the driving industry, you are still protected by the buffer zone called regulation. If you take your job seriously and you belong to these industries, it will be worth staying up-to-date with current technological advances and especially regulations. The first introduction of automation technologies always aids the worker before phasing them out completely. Truck drivers already benefit from cruise control, for example. Driverless trucks may still need operators on board to oversee functionality and extreme edge cases the computer cannot resolve. What happens if a crazy person jumps in front of the truck to mess with it? They are more likely to do this if the truck is completely uncrewed. The same goes for potential hacking scenarios. Thankfully for you, driverless cars are a regulatory nightmare and probably won't see a mass adoption for a few decades.

Question: I've heard that AI can drastically change the world, bring about world peace and even end poverty. Is there any truth to this?

Answer:

It really depends on whom you ask, as nobody understands what AI will look like in 30 or 50 years down the line. Artificial general intelligence that converges with humanities goals will be able to perform many tasks better than we do. It can then work on these problems without needing to eat or sleep, bringing about advances at an exponential rate. Such systems could probably solve the most pressing problems today, like climate change and food logistics. Some believe they will bring about a new age of abundance where people are not required to work anymore. The wealth created by this intelligence will fuel a universal basic income. Those are the most optimistic projections. A more level head explanation may be a world where general AI is limited to select corporations and governments, further contributing to wealth inequality. Still, another practical explanation is that general AI is a fad idea that will never see reality.

Question: How worried should I be about AI in general?

Answer:

For the average person, there is little to worry about. Most applications of machine learning are completely benign. There are some privacy and ethical concerns, but then again, most people will not be affected by them. Even if the military adopts these weapon systems, there is an asymptotic risk that they will be used against you. Which is to say, virtually zero. Losing a job may be a concern, in which case you have a long head start to think about what other industry you can switch into. Remember for every terrible application of the technology there is a good potential benefit to humankind. Imagine a world were graphics content is removed from family websites like Facebook. Imagine a world where driving is as safe as flying a plane, where young drivers are not at an increased risk for fatalities. Imagine a world where market inefficiencies are smoothed out, leading to an increase in the economy. Imagine a world where cancer is no longer a major killer due to early detection and advanced drug discovery. Imagine a world where the elderly live longer and are cared for with the dignity and respect that they deserve, no longer having to worry about abusive caretakers or the embarrassment of loss of autonomy.

Question: If general AI is discovered and it does go bad, what can humans do?

Answer:

Like with most intractable scenarios, prevention is the cure. Trust that companies are doing the right thing to ensure that their systems do not run off course. In the future, AI legislation will be a common topic of debate. If you are not currently a voter, you may reconsider when the fate of humanity is at stake. There is a striking hypothesis about the creation of general AI that posits that the sooner it is discovered, the less of an existential threat it will be for humanity. This is because the longer it takes, the better technology and systems it will have at its disposal. Imagine if nanotechnology 3D printing machines are in active use when the general AI comes online. If it were to go rogue, it could very quickly begin transforming the face of the Earth into a massive computer.

Conclusion

Thank you for making it through to the end of *Artificial Intelligence: What You Need to Know About Machine Learning, Robotics, Deep Learning, Recommender Systems, The Internet of Things, Neural Networks, Reinforcement Learning, and Our Future*. It should have given you a better understanding of AI. The hope is that you can understand internet headlines that talk about AI without feeling completely clueless. The knowledge contained in this book should have been enough for that end, and to encourage making your own conclusions.

The field of AI is vast and ever-changing. Many methods of AI like genetic algorithms, statistical inference, and stochastic processes were not covered in this book, but there are countless resources out there that do. Additionally, you can find a complete history of AI on the Wikipedia page with current developments as well. If you wish to pursue artificial intelligence as a course of study or even as a career, be prepared for lots of math and computer programming.

Finally, if you found this book useful in any way, a review on Amazon is always appreciated!