FREEBOOK

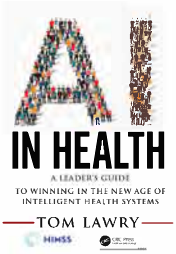
HIMSS Resource: AI in Healthcare

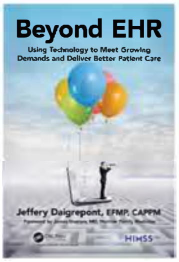
Routledge
Taylor & Francis Group

HIMSS

www.routledge.com

Contents





1. **AI is One Thing and Many Things**
Tom Lawry
AI in Health

2. **Using Artificial Intelligence in Healthcare**
Jeffery Daigrepont
Beyond EHR

3. **Applied Machine Learning for Healthcare**
Prashant Natarajan, Bob Rogers
Demystifying Big Data and Machine Learning for Healthcare

All HIMSS members receive 25% off through the code provided to you by member services.** Visit the **HIMSS Bookstore** to view all of the latest HIMSS titles from CRC Press.

CRC Press, Taylor & Francis Group is proud to be the exclusive distributor and publisher for books published by HIMSS, a global advisor and thought leader supporting the transformation of the health ecosystem through information and technology.

At CRC Press, we understand that health information and technology and informatics are growing topics of increasing complexity in the healthcare sector. We will continue to bring you the latest content in healthcare management, informatics, artificial intelligence, patient experience, and information management and systems from HIMSS as it is published.

***Valid for online orders, excludes eBooks.*

Chapter 3

AI Is One Thing and Many Things

Alan Turing was an English mathematician, computer scientist and theoretical biologist who is considered to be the father of modern computer science. He formed the concept of algorithms and came up with the idea of a machine that was capable of computing anything that could be computed. His seminal work is the basis for the 2014 movie *The Imitation Game*. In this film, Turing is seen using an early form of computer science to crack the secret code of the Germans. His work saved lives and helped to end World War II.

And while Turing first put forward the notion of a universal machine capable of developing and using algorithms in the 1930s, it wasn't until the summer of 1956 that the world was first introduced to the term "AI." It was a situation that, should a movie be made, would more likely resemble a geeky version of *Animal House*.

John McCarthy was a young assistant professor of mathematics at Dartmouth College, who decided to hit up the Rockefeller Foundation for a grant to bring together mathematicians and scientists to explore emerging and fringe concepts of imbuing machines with "intelligence." With grant in hand, McCarthy invited a few dozen of his friends and top minds in math to a 2-month rolling nerd-party. The group checked in to the red-bricked Hanover Inn then sequestered itself on the top floor of the Dartmouth Math Department building. They would spend the summer in a continuous brainstorming session, which led to many of the concepts still used today.

From this gathering came the term "AI" along with many concepts of how machines could use language, form abstractions, and begin to solve problems previously reserved for humans. Many of the precepts formulated in the summer of 1956 remain relevant today. This includes concepts that formed the basis for things like Machine Learning (ML) and Natural Language Processing (NLP). In many regards, the rolling nature of what was started by Turing in the 1940s and McCarthy in the 1950s continues as we evolve in the capabilities and use of AI today (We'll learn more about the history of AI and its impact today in the next chapter).

AI Is One Thing and Many Things

The terms "Artificial Intelligence" and "AI" are often bandied about as if there is a singular definition that is universally understood. In reality, AI is not one technology. A simple Google search of "what is AI?" turns up 4.3 billion results.

AI is an umbrella term that includes multiple concepts and technologies used individually and in combination to add intelligence to computers and machines. For example, it is commonplace for the terms AI and Machine Learning (ML) to be used interchangeably. In reality, they are not the same. This misperception often causes confusion (as you will learn in this chapter, ML is usually classified as a subset of AI).

Equally important to understand is that the set of capabilities that fits within the term AI is evolving. As a health leader it is not imperative to be an expert in AI. It is, however, useful to understand the basics on which intelligent systems are built. To effectively plan and lead your organization's AI strategy it's important to have a general understanding of what it is. This includes having a framework for how the "components" and "capabilities" that exist can be used in defining, building, and managing intelligent systems. It is not necessary to know how AI actually works (though such knowledge never hurts). What is important to understand is what capabilities exist and how to put them to work in service of your organization's mission and goals.

When it comes to defining AI, there is generally broad agreement on what it is, but little or no standardization when it comes to organizing the components or building blocks of AI into a universally accepted taxonomy.

With this in mind, this chapter provides a working definition of AI as well as a nontechnical framework for understanding and applying these "AI building blocks" to your AI strategy. This book, nor this chapter, are not designed as a technical resource for AI. For those interested in going deeper on technical definitions and capabilities there are many great resources available.

Defining AI – General

Let's start with a simple but functional definition of AI:

> Artificial Intelligence (AI) is an area of computer science that emphasizes the creation of machines that work and react like humans. This means systems that have the ability to depict or mimic human brain functions including learning, speech (recognition and generation), problem-solving, vision and knowledge generation.

AI is a constellation of technologies that allow computers and machines to sense, comprehend, act, and learn. Unlike IT systems of the past that merely generated or stored data, the value of AI systems is that they can increasingly learn, adapt, and complete tasks in ways that are similar to a human being. In this regard AI imbues machines with intelligence. To understand and test this concept one can turn to what was posited by Alan Turing more than 60 years ago as a simple but effective definition of whether a machine is intelligent. The "Turing test" states that a machine can be considered "intelligent" if a human cannot distinguish the responses of a machine compared to responses from a human.[1]

AI Building Blocks

The general definition offered above is a descriptor for what AI *is*. The components of AI described below is the "how" part of the equation. These are the functional capabilities provided by AI today.

For the sake of providing a framework for understanding AI capabilities, let's take the broad definition noted above and break it down further into AI building blocks. A building block is an explicit component of AI that mimics a capability found in humans. We'll further separate the definition into two types of building blocks: Machine Learning and Cognitive Services.

Machine Learning (ML)

Ask anyone today what type of AI project they are working on and the most likely answer will be something they want to predict. The ability to predict things comes from **machine learning (ML),** which is a subset of AI. ML provides software, machines, devices, and robots with the ability to learn without human intervention or assistance or static program instructions.

Machine learning evolved from the study of pattern recognition and computational learning theory. The term was coined by AI pioneer J.A.N. Lee in 1959, who defined it as a "field of study that gives computers the ability to learn without being explicitly programmed.[2]"

Interestingly, Samuel is best known for his work with AI and computer gaming. If you think your kids spend too much time gaming, you can thank Samuel as he is credited with creating the world's first computer game. Known as the "Samuel Checkers-playing Program", it was among the world's first successful self-learning programs that demonstrated the fundamental concept of AI. Other things that were to come decades later, such as use of IBM Watson to be the best players at chess or the game show Jeopardy, have their origins rooted in Samuel's groundbreaking work.

Today, machine learning is at the forefront of making AI real in healthcare. It's at the heart of our ability to predict things we care about. It's used to identify the root cause of quality problems, recommend treatment options to clinicians, drive smartphone apps for consumers, improve operational efficiencies, and more.

Machine learning-enabled processes rely on the development and use of computerized algorithms that "learn" from data sets rather than strictly following rule-based, preprogrammed logic. An **algorithm** is a mathematical model based on sample data, known as "training data."

Machine learning uses algorithms to identify patterns in the data and then make predictions from those patterns with a degree of certainty. Based on input data, machine-learning can improve its accuracy over time through a feedback loop and modify the approach it takes in the future—hence the term "learning."

It's important to understand that the "learning" part of "machine learning" is purely mathematical and has little to do with understanding what the algorithm has learned. This is different than when humans analyze data where we build an actual understanding of the data to a certain extent. We'll go deeper into understanding and applying the differences in what AI is good at versus humans in Chapter 5.

In spite of lacking deliberate understanding and being a mathematical process, machine learning can prove useful in many tasks. It provides many AI applications the power to mimic rational thinking, given a certain context when learning occurs by using the right data.

Within the general category of machine learning are various models used to create different types of algorithms. For example, supervised learning algorithms are a type of machine learning that involves direct human supervision and use labeled data to predict future outcomes after being trained based on past data. Unsupervised learning focuses on the use of data in an unguided fashion that has not been labeled, classified or categorized to complete a cluster analysis that looks for relevant patterns or trends within the data.

Reinforcement learning is the training of machine learning models to make a sequence of decisions by identifying patterns and making decisions with minimal human intervention. It focuses on developing a self-sustained system that, throughout contiguous sequences of trials and errors, improves itself based on the combination of labeled data and interactions with new data.

Even within these three models are many terms you may hear about. For example, neural networks are a form of supervised learning consisting of interconnected units (like neurons) that process information by responding to external inputs, relaying information between each unit. The process requires multiple passes at the data to find connections and derive meaning from undefined data.

Deep learning uses large neural networks with many layers of processing units, taking advantage of advances in computing power and improved training techniques to learn complex patterns in large amounts of data. Common applications include image and speech recognition.

With the use of ML, AI is increasingly good at being able to sense and predict things we care about. Like which patients are at high risk of readmissions, falls, or unexpected deterioration. It can help with predictions of which treatments may produce the best outcomes. It's already making diagnostic images more "intelligent."

There are many outstanding technical books and resources available to describe the various types and forms machine learning algorithms can take (linear regression, logic regression, decision tree, and decision forests) as well as the various languages and methodologies used to create them (R, Python, Lisp, to name a few). Unless you are a budding data scientist, it is sufficient to recognize that there are various types of machine learning models and algorithms. This diversity of the languages and models allows data scientists and developers to design an approach using various models and languages that best fit the type of problem to be solved.

EARLY INTERVENTION FOR PATIENTS AFFECTED BY COPD USING AI AND ML

Medical experts at one of the largest health boards in Europe are using ML to remotely monitor and treat people with life-threatening lung conditions in their own homes.

NHS Greater Glasgow and Clyde are using ML to predict pulmonary risk with the goals of improving care for patients suffering from Chronic Obstructive Pulmonary Disease (COPD). COPD is a group of lung conditions that cause breathing difficulties. The World Health Organization (WHO) forecasts COPD to become the third leading cause of death worldwide by 2030. Within the United Kingdom, COPD affects 1.2 million people and is the second most common cause of emergency hospital admissions and accounts for one in eight of all UK hospital admissions. As a result, it is also a key driver in rising NHS costs with each unplanned hospital trip costing about £6,000.[16]

The project is designed to improve quality and reduce emergency hospital admissions among high-risk COPD patients through the use of AI to help recognize patterns in a patient's condition and provide early warnings to aid intervention and prevention. It does this by remotely monitoring patients' symptoms, physiology, and treatment in the home with ML algorithms being used to detect or predict clinical deteriorations such as acute exacerbations of COPD.

Beyond monitoring and managing the health of individual citizens, the initiative is also designed to enable hospitals across Scotland to predict patterns in COPD admissions and estimate their length of stay, helping them to better manage resources. AI-guided predictive modeling and cloud-based decision support dashboards will enable focused patient–clinician communication, target clinical team's resources, and facilitate early activation of self-management and tiered COPD treatment interventions.

The project aligns closely with the ambitions set out in the recently published Scottish Government Digital Health and Care Strategy by supporting the shift of care into the community, reducing readmission rates, and empowering citizens to more effectively manage their own care.[17]

The COPD project initially involves 400 patients. The goal of better prediction and management of this initial trial group of patients is to improve health status, have patients more deeply engaged in monitoring and managing their health while reducing hospitalizations which equates to a potential cost saving of £1.2 million per year. The service also aims to reduce emergency hospital admissions per year. If this is scaled across 20% of the highest risk COPD patients, the potential NHS savings equates to £1.4 billion.[18]

Cognitive Services

Beyond the power of making predictions through machine learning, there are a growing number of applications or solutions that can be categorized as cognitive services. As the title implies these AI building blocks mimic specific human functions including perception (e.g., seeing, hearing), language, thinking, and learning.

Such functionality is available today and most often deployed through the use of an **"Application Programming Interface or API."** An API is a preset group of computer commands and protocols used by programmers to create software or interact with external systems. APIs provide developers with the ability to efficiently perform common operations without writing the code from scratch.

While the list of Cognitive Services is constantly evolving, here are the most common types of applications.

Computer Vision: It is a field of computer science that works on enabling computers to see, identify, and process images in the same way that human vision does. It then provides appropriate output to complete a task. Computer vision is a form of AI, as the computer must interpret what it sees and then perform appropriate analysis.

Vision services allow humans to gain insights from images, pictures, and video. These capabilities range from detecting faces in pictures, automated image analysis, text, or video moderation as well as person recognition. The goal of computer vision is not only to see but also process and provide useful results based on the observation. For example, the application of computer vision in healthcare ranges from detecting faces to aid in things like member verification and patient registration to clinical applications including autodetecting abnormalities in diagnostic and pathology images.

A fun example of computer vision at work is the website www.how-old.net, which allows users to select or upload an image of a person to then have computer vision and machine learning estimate the age of the person in the image.

COMPUTER VISION TO SCREEN FOR HAZARDOUS SKIN LESIONS

Vision intelligence is being used to help screen for skin lesions that are potentially cancerous. Today various smartphone apps are available to download that allow users to take a picture of a skin lesion to assess the likelihood of being a cancer concern.

In a recent study by the International Skin Imaging Collaboration (ISIC) and the MedUni Vienna, researchers also found that when it comes to the diagnosis of pigmented skin lesions, AI is better than humans. This database includes benign (moles, sun spots, senile warts, angiomas, and dermatofibromas) and malignant pigmented lesions (melanomas and other types of skin cancer).

When it came to comparing "man versus machine" results, researchers found that the best humans diagnosed 18.8 out of 30 cases correctly, while the best AI performance diagnosed 25.4 out of 30 cases correctly. The team says two-thirds of all participating AI algorithms were better than humans, and this result had been evident in similar trials during the past year.

They also mention that although the algorithms were superior in this experiment, this does not mean that the machines will replace humans in the diagnosis of skin cancer as there are many other factors beyond initial diagnosis.[22]

Knowledge Extraction: It allows for the identification, organization, and extraction of specific information and knowledge from large amounts of preexisting data and information. As the amount of data and information increases in healthcare the ability to extract and mine this data to acquire new knowledge becomes vitally important.

Knowledge extraction allows us to use massive quantities of data and information to look for patterns that humans simply don't have the ability or time to see.

To illustrate this point, researchers from the Lawrence Berkeley National Laboratory used machine learning to reveal new scientific knowledge hidden in old research papers. Using just the language in millions of old scientific papers, a deep learning algorithm was able to make new scientific discoveries by sifting through scientific papers for connections humans had missed. And while the experiment was focused on new discoveries in material science the process could just as easily be applied to other disciplines such as medical research and drug discovery.[3]

Speech: The speech component of AI is getting a lot of uptake in healthcare today as it provides the ability to implement speech translation and recognition features into applications and workflows to make an automated process more human (understand what a human is saying). This area also provides the ability to convert text to speech and vice versa on-the-go to understand user intent and interact with patients and consumers.

Speech recognition has been around a while, but the AI-enhanced capabilities have brought the capabilities of machines to be on par with the speech and language capabilities of humans.

Language Understanding: Allows a computer application to understand what a person is saying and wants in their own words.

Natural Language Processing (NLP): NLP enables computers to derive computable and actionable data from text, especially when text is recorded in the form of natural human language (i.e., phrases, sentences, paragraphs). This technology allows humans to record information in the most natural method of human communication (narrative text) and then enables computers to extract actionable information from that text. NLP is also capable of analyzing the often-nonstandard grammatical constructions common in medical language. Natural language understanding (NLU) is a subset of NLP that uses reasoning, inference, and semantic searching to help clinicians make decisions and take action.[4]

Text Analytics: Provides NLP over raw text for sentiment analysis, key phrase extraction, and language detection.

Search: Search is one of the most important services for nearly every application or solution nowadays. In order to implement a search service, it is essential to provide the best possible results.

CHATBOTS – BRINGING CONVERSATIONAL AI TO HEALTHCARE

Chatbots are part of a growing trend of intelligent virtual assistants in health which are improving customer service, increasing patient engagement, automating repetitive tasks, and allowing staff to spend more time in higher-value customer service activities. From a friendly voice to a text box flashing a simple "how can I help," intelligent assistance comes in the form of smart speech and text applications.

According to Gartner, 25% of customer service and support operations are now integrating virtual customer assistant (VCA) or chatbot technology across all engagement channels, up from less than 2% in 2017.[11]

As more customers engage with health providers through digital channels, VCAs and chatbots are being implemented for handling customer requests on websites, mobile apps, consumer messaging apps, and social networks. All of this is being driven by improvements in NLP, ML, and intent-matching capabilities.

Organizations using these intelligent solutions report a reduction of up to 70% in call, chat, and/or email inquiries after implementing a VCA or chatbot, according to Gartner research. They also report increased customer satisfaction and a 33% saving per voice engagement.[12]

Today, healthcare chatbots are a mix of both patient-only and patient–clinician applications that connect the two groups: administrative and diagnosis and treatment-related activities.

In the world of health plans, consumers often don't want to wade through a website or app, and really don't want to make a call and hear, "Please continue to hold." They just want to type a question and get an answer.

To the rescue is a chatbot created by Premera Blue Cross, the largest health plan in the Pacific Northwest of the United States. Premera' s virtual assistant is called Scout and uses easy-to-understand, text-based chat to help customers quickly know where to get information on claims, benefits, and other services. Users can talk to it in a natural way and get answers back in a way consumers actually speaks. With an avatar of a serene blue owl, Premera Scout helps customers "self-serve" basic questions at any hour, while giving customer-service employees more time to handle complicated requests.[13]

Premera Scout is helping deliver better customer service to the health plan's 2 million customers in Washington and Alaska.

Fueled by advances in AI and the boom in messaging apps, chatbots take advantage of AI building blocks including speech, text, and natural language processing capabilities to streamline what is often a complex maze of telephone prompts and long wait times for talking to a human.

Another example of smart customer service is Quest Diagnostics, a comprehensive clinical laboratory. As a Fortune 500 company Quest Diagnostics has about 45,000 employees and operates in the United States, United Kingdom, Mexico, and Brazil. Every year, tens of millions of adults are asked to contact Quest Diagnostics for healthcare-related services that range from routine blood work to complex genetic and molecular testing. In today's increasingly self-service healthcare industry, details such as where to go, when, and what to do beforehand are typically up to the patients to figure out for themselves.

The Quest chatbot helps people who visit the Quest Diagnostics website during call center hours find testing locations, schedule appointments, and get answers to nonmedical questions such as whether to fast before a blood draw or when to expect results. If the bot is unable to answer a question or the user gets frustrated, the bot will transfer the user, along with the context of the conversation, to a person who can help – all without having the user pick up the phone. Based on a user-experience survey, Quest found that about 50% of their clients would prefer to engage with a chatbot instead of a search box or the frequently asked questions feature on a website.[14]

Chatbots and conversational AI is creating a new form of patient engagement and customer service that can be personalized to the needs of the consumer while, as needed or requested, leveraging the skills and availability of staff to get involved to allow the experience to remain human-focused.

Here are some of the top areas where conversational AI is driving innovation and an improved consumer and patient experience.

Repetitive Customer Service Inquiries and Interactions: Chatbots are now supporting and supplementing staff in answering routine questions around registration, scheduling, benefits, and billing.

Patient Engagement and Care Coordination: Increasingly chatbots are being used in issuing reminders, scheduling appointments, and automating prescription refill requests.

The benefit of chatbots in patient engagement is the ability to provide advice and information for routine things like healthy lifestyle tips or specific reminders for those with specific conditions like diabetes (all of course under the direction of the clinician driving the care plan). Another example of its benefit in patient engagement is the ability to provide intelligent reminders for things like helping patients with when to take their pills.

For example, in the UK, a chatbot can be used to search volumes of health information from the National Health Service (NHS). This includes choices for healthcare advice and resources that support consumers reviewing conditions and treatment options.[15]

Chatbots are increasing in sophistication, which allows users to enter their symptoms via chat, view a list of related conditions, and through a series of prompts identify a patient's potential condition.

Triage: Whether it's getting a consumer to the right person to answer a question or helping in the gathering of information to get to the root of a patient medical problem, AI-powered chatbots are creating a more efficient process for clinicians and consumers.

The goal of using chatbots and virtual assistant is to help clinical and operational staff optimize their time while putting more information in the hands of patients and consumers. In the end, this intelligent self-service option helps to drive better outcomes and reduce costs.

Applying AI Building Blocks

The AI "building blocks" described above are often used in combination and deployed in a variety of ways with other technologies (like sensors) to drive value in automating or augmenting work heretofore done by humans.

AI is sometimes classified by the level of sophistication or type of use. For example, **Narrow AI** (Artificial Narrow Intelligence or ANI) is good at performing a single task, such as predicting which patient is likely to be a no-show for an appointment. AI components like machine learning, computer vision, and natural language processing are currently in this stage. As such narrow AI is a bit like a digital idiot savant, it excels at one particular type of task within a limited context but isn't able to take on tasks beyond what it was designed to do. Even when pushing the boundaries of today's AI, most everything being done is through Narrow AI. For example, self-driving car technology is still considered ANI, or more precisely, a coordination of several narrow AIs[5].

Understanding that almost all AI deployed today is considered "narrow AI" is something we'll build on in the next chapter, as creating value is heavily dependent on understanding how to leverage and balance the capabilities of AI with the unique capabilities of your human teammates.

General AI (also known as Artificial General AI or AGI) is the type of AI that can understand and reason across its environment as a human would. General AI has always been elusive. This category of AI is where many organizations aspire to be someday (e.g. IBM's Watson), but any true form of this is not likely on the short-term horizon.

As you will learn in the next chapter, humans might not be able to process data as fast as computers, but they can think abstractly and plan, and can solve problems based on their experience and creativity. These factors, which are vitally important in delivering health services, are not found in the realm of what computers can replicate today nor likely anytime soon.

Finally, there is another category known as **Super Intelligence** (Artificial Super Intelligence or ASI) that is a level of computer sophistication where machines become smarter than the humans that create them. This is the stuff that becomes fodder for science fiction movies. It's also the type of AI you occasionally here people like Elon Musk waxing over as they ponder the dangers to mankind should we reach this level of AI capability.

For now, recognize that pretty much everything being done today falls into the Narrow AI category. Investments that are being made in the tech industry are designed to move systems closer to the General AI category.[6]

Here are some of the common ways AI is packaged and deployed today:

AI Apps: Web or mobile applications are infused with AI capabilities, such as vision, language, or ML. For example, AI is pervasive in our daily lives. From anonymized data from smartphones and other data, AI analyzes the speed and movement of traffic at any given time to predict when you will reach your destination and the best route to do so. As you make an online transaction with your credit card, AI is running in the background to monitor and predict whether the charge is fraudulent. When it comes to health, thousands of consumer health apps help monitor body functions, provide alerts for various health indicators, or guided recommendations on everything from nutrition to maintaining emotional and mental wellness.

Bots and Conversational AI: A bot is an automated application used to perform simple and repetitive tasks that are often time consuming for a human to perform. **Conversational AI** makes use of speech and language building blocks to automate communications and create personalized customer experiences that are scalable.

With consumers increasingly looking to access information on demand, the use of bots and conversational AI provide short and high value interactions with customer and staff through task automation and automated workflows. The goal in the use of Bots and Conversational AI is to improve the customer experience while reducing the need for lower-value human interactions.

In one survey of consumers nearly 70% saw chatbots as the best way to get instant answers to their questions and over one-fifth (21%) saw chatbots as the easiest way to contact a company.[7]

Intelligent IoT: The Internet of Things (IoT) is a network of Internet-connected devices that communicate embedded sensor data to the cloud for centralized processing. These sensors can be embedded in everyday items such as cell phones, digital weight scales, or wearable health and medical devices; or they could be components of larger machines and systems such as medical imaging or lab systems.

The introduction of intelligence with IoT enables health organizations to reimagine existing services or create new types of services that cut across historical care settings. Intelligent IoT also allows for the ability to improve operational efficiencies in areas that include smart remote patient monitoring or improved predictive maintenance of equipment and facilities.

IoT IN ACTION: AN ECG ON YOUR WRIST

There's a new breed of wearables emerging making it easier for people to continuously monitor their hearts with medical-grade devices.

While fitness trackers are popular among the "quantified-self" crowd, they aren't serious medical devices. Variables such as an intense workout or a loose band can affect the sensor reading your pulse. But an electrocardiogram (ECG), the kind doctors use to diagnose abnormalities before they cause a stroke or heart attack, requires a visit to a clinic, and people often fail to take the test in time.

ECG-enabled smart watches, made possible by new regulations and innovations in hardware and software, offer the convenience of a wearable device with something closer to the precision of a medical one. While standard fitness trackers typically employ a single sensor, a real ECG has 12. And no wearable can yet detect a heart attack as it's happening.

A number of IoT wearable devices are now receiving clearance for use that can detect things like atrial fibrillation (AFib), a frequent cause of blood clots and stroke. Worldwide more than 33.5 million people have AFib.[19] In the United States more than 750,000 hospitalizations occur each year because of AFib and contribute to nearly 130,000 deaths each year.[20]

At the moment, even people at risk for AFib with the best access to care get only two or three ECGs a year. Preventive screening through wearables could if widely implemented, potentially save thousands of lives. In addition to the monitoring work underway is a move to use deep learning algorithms with the data coming from wearables and look for new ways of using it. Apple and Johnson and Johnson partnered recently to a study that screens for stroke risk. And AliveCor's software developed in conjunction with the Mayo Clinic has been granted accelerated clearance to use deep learning on ECGs to screen for hyperkalemia, (elevated potassium levels in the blood) that puts people with kidney disease at risk for arrhythmia and death.[21]

Intelligent Robots: According to a study by Accenture, robot-assisted surgery is estimated to produce $40 billion in near-term value to health organizations.[8] With the help of AI, robots can use data from past operations to guide surgeons to improve existing surgical techniques and reduce the invasive nature of some surgeries. One study of the use of smart robotics in orthopedic surgeries resulted in five times fewer complications.[9]

Looking ahead, heart surgeons are now being assisted by a miniature mobile robot known as HeartLander, which facilitates minimally invasive therapy to the surface of a beating heart. Under the control of a physician, the **robot** enters the chest through an incision below the sternum and then autonomously navigates to the specified location on the heart to administer the therapy.[10]

Notes

1 Stuart J. Russell, Peter Norvig. *Artificial Intelligence: A Modern Approach* (3rd ed.). Upper Saddle River, NJ: Prentice Hall. ISBN 978-0-13-604259-4, 2009.
2 J.A.N. Lee, Arthur Lee Samuel, Computer Pioneers, IEEE Computer Society, https://history.computer.org/pioneers/samuel.html.
3 Vahe Tshitoyan, Unsupervised word embeddings capture latent knowledge from materials science literature, *Nature*, 2019, www.nature.com/articles/s41586-019-1335-8.epdf?referrer_access_token=gjs4tfb7-T50BFnuqtYx5NRgN0jAjWel9jnR3ZoTv0P9QxlcO86f_GXZRxwYijrqa11Mx55SgniZXv55YKOR_sn816NK2x0O46Vim16XrS-SjyP9GMXeDQinUN75ES6enlxK__J5UabR6JdgR19bZSVLL5ZsK8146qMcipEbItW65C8aSk29Q_BfrKz4Gb5-kjz3m7dIaoRxs3e1I6qW4022QZ6aZMaOPxlATK7OOqj8lrhj-yufvROMPdStMZjAEK-efja6SfW5n-6xhZuV3zQTFR_u132mC6hkt8Zqp29_su0pmsC0jrneuemHnqg8&tracking_referrer=www.vice.com.
4 Laura Bryan, MS, CHDS, AHDI-F, With Natural Language Processing You Can Have Your Cake and Eat It Too, HIMSS, 2019, www.himss.org/news/natural-language-processing-you-can-have-your-cake-and-eat-it-too.
5 Ben Dickson, What is Narrow, General and Super Artificial Intelligence, TEchTalks, 2017, https://bdtechtalks.com/2017/05/12/what-is-narrow-general-and-super-artificial-intelligence/.
6 Dean Van Nguyen, Elon Musk Calls Artificial Intelligence 'Our Biggest Existential Threat', Silicon Republic, 2014, www.siliconrepublic.com/machines/elon-musk-calls-artificial-intelligence-our-biggest-existential-threat.
7 2017 Chatbot Survey, Ubisend, 2017, www.ubisend.com/insights/2017-chatbot-report.
8 Matt Collier, Artificial Intelligences-Healthcare's New Nervous system, Accenture, 2017.
9 Brian Kalis, Matt Collier, Richard Fu, 10 Promising AI Applications in Health Care, *Harvard Business Review*, 2018, https://hbr.org/2018/05/10-promising-ai-applications-in-health-care.

10 HeartLander, Carnegie Mellon University Robotics Institute, 2016, www.ri.cmu.edu/project/heartlander/.

11 Gartner Says 25 Percent of Customer Service Operations Will Use Virtual Customer Assistants by 2020, Gartner Inc. 2018, www.gartner.com/en/newsroom/press-releases/2018-02-19-gartner-says-25-percent-of-customer-service-operations-will-use-virtual-customer-assistants-by-2020.

12 Ibid.

13 Vanessa Ho, Skip the Hold Music and Meet Premera Scout, a Chatbot That Helps You with Health Care, Microsoft Blog, 2018, https://news.microsoft.com/transform/premera-scout-chatbot-helps-with-health-care/.

14 John Roach, Microsoft Service Helps Healthcare Organizations Develop and Deploy Virtual Health Assistants, Microsoft Blog, 2019, https://blogs.microsoft.com/ai/microsoft-healthcare-bot-service/.

15 Kumba Sennaar, Chatbots for Healthcare – Comparing 5 Current Applications, Emerj.com, 2019, https://emerj.com/ai-application-comparisons/chatbots-for-healthcare-comparison/.

16 NHS Greater Glasgow & Clyde is Monitoring COPD Patients with Wearables, NSTech, 2019, https://tech.newstatesman.com/cloud/nhs-greater-glasgow-clyde.

17 Scotland's Digital Health and Care Strategy, The Scottish Government, 2018, www.digihealthcare.scot/wp-content/uploads/2018/04/25-April-2018-SCOTLANDS-DIGITAL-HEALTH-AND-CARE-STRATEGY-published.pdf.

18 Doctors Are Using Technology to Help Patients with a Serious Lung Disease Get Treatment in Their Homes, Microsoft News Centre UK, 2019, https://news.microsoft.com/en-gb/2019/03/26/doctors-are-using-technology-to-help-patients-with-a-serious-lung-disease-get-treatment-in-their-homes/.

19 Atrial Fibrillation: Facts, Statistics, and You, Healthline, 2018, www.healthline.com/health/living-with-atrial-fibrillation/facts-statistics-infographic.

20 Ibid.

21 Don Hon, Wristwatch Heart Monitors Might Save Your Life—and Change Medicine, Too, *MIT Technology Review*, 2019, www.technologyreview.com/s/612929/wristwatch-heart-monitors-might-save-your-lifeand-change-medicine-too/.

22 Phillip Tschandi PhD et al., Comparison of the accuracy of the human readers versus machine-learning algorithms for pigmented skin lesion classification: An open, web-based, international, diagnostic study, *The Lancet Oncology*, 2018, www.thelancet.com/journals/lanonc/article/PIIS1470-2045(19)30333-X/fulltext#.

Chapter 5

Using Artificial Intelligence in Healthcare

The adoption of artificial intelligence (AI) in healthcare is rising and delivering positive signs in assisting and solving a variety of problems for patients, providers and hospitals. *Forbes* Magazine reports a 14× increase in AI startups since 2000. With investment in the industry up six-fold, topping out at over $3 billion, spending on AI is likely to surpass all the outlays on EHRs combined. This chapter explores the current uses of AI, e.g., listening devices and Web applications, facial recognition and its use in the exam room, and ethical questions that arise from the use of AI in healthcare. Lastly, we consider how to establish a strategy for implementing AI in your own practice or hospital.

What Is Artificial Intelligence?

First, let's define artificial intelligence.

- The essential requirement of AI is intelligence, defined as the ability to acquire and apply knowledge and skills. It is the capacity to interact (speech, vision, motion, manipulation), reason, learn, adapt and think abstractly, as measured by objective criteria, such as test taking. Also, the AI must be capable of adapting to the outcomes or variables on its own.
- The term *artificial intelligence* is an umbrella term for machines capable of perception, logic and learning. Today, there are two types of AI:

- **Machine learning** is a form of AI that employs algorithms capable of learning from data to make predictions or decisions; as the machine's exposure to data increases, the performance capabilities are improved. In some cases, this may be a simple "if, and, do what" programming logic. For example, the "if-then statement" is set as the most basic of all the programing controls. It tells your program to execute a certain code only if a statement is true. For example, the heart monitor will send an alert **if** blood pressure drops below a certain threshold, but only **if** the program detects that the monitor is connected to a patient. An example of the program language could look like the following
  - Alert Monitor Alarm () {
    - // the "if" clause: *blood pressure dropping*
    - if (connected to patient){
    - // the "then" clause: *sound the alarm*

      A single AI program could have millions of IF-AND-DO-WHAT programing lines of code/scripts for making split-second decisions like how the human brain might work. For example, if you see a person running toward you, your reaction will vary considerably depending on data already stored in your brain. Do you know the person? Do they look excited to see you? Are they running up to you to give you a hug? **If** all are true, the **do what** outcome would likely be to stay calm. If the person running toward you looks unfamiliar, angry and is carrying a knife, the **do what** outcome is to RUN!!!! Should such an event occur, the brain (or AI) quickly responds and adapts based on the incoming data. (See Figure 5.1 for an example of machine learning logic.)
- **Deep Learning AI** is also based on the IF-AND-DO-WHAT programing concept but will act more like the human brain and adjust behavior based on prior outcomes. Using the same example of a person running toward you, the brain will record a massive amount of data from that event (e.g., location, time, images, facial recognition, etc.), then use this stored information to make better decisions in the future, such as avoiding dark alleys at 2:00 AM. Deep learning uses many-layered neural networks (computer systems based on the human brain and nervous system) to build algorithms that find the most efficient way to perform a task based on vast sets of data. Deep learning will typically improve over time by adding all

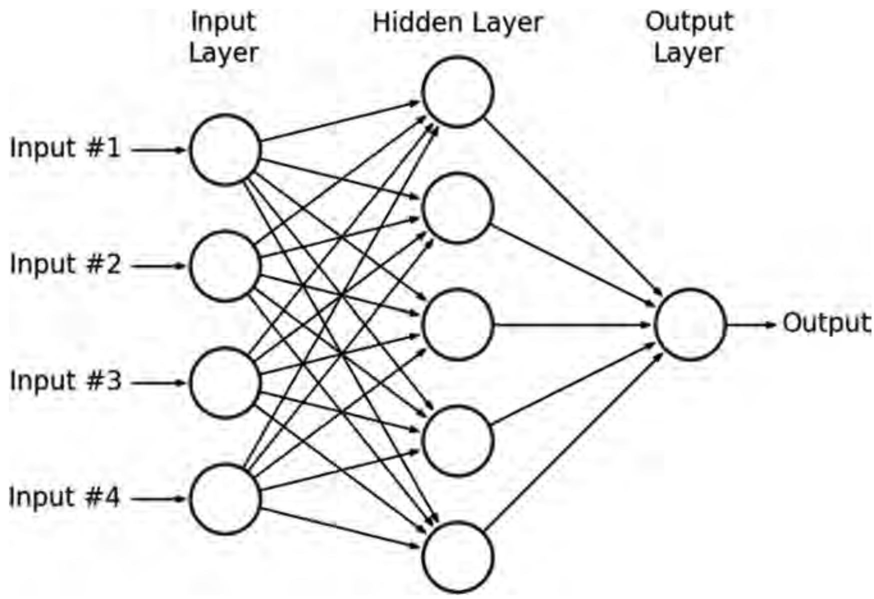


Figure 5.1 A hypothetical example of multilayer perceptron network.

past outcomes to the logic for future decisions. (See Figure 5.2.) One major limitation of the deep learning method is the inability to apply a conscience to decision-making. A major concern with the deep learning method is the possibility for bias and prejudice to seep into the algorithms. Bias in AI happens when the data used are unrepresentative of reality, or reflect the existing prejudices of the developer/programmer. An example of this was recently seen in AI software used to help judges in sentencing criminals based on the probability of the person being a repeat offender. Although AI is becoming better every day, the algorithms we see every day still have a long way to go before being safely applied to the criminal justice system.

While we do not foresee AI replacing all humans, the study of the ethics and risks of machine involvement in patient care, compared to traditional methods, has yet to catch up to the technology. More concerning is the overall impact on our society and the shifts in inequalities that AI is expected to cause. Specifically, AI is expected to eliminate 40% of all repetitive jobs over the next 20 years. Examples include call centers, patient check-in, registration, triage, collections, accounts receivable (AR) follow-up and campus delivery services. In March of 2019, UPS launched a new service using drones to transport blood and other medical supplies between the various buildings at the WakeMed Health and Hospitals medical campus in Raleigh,

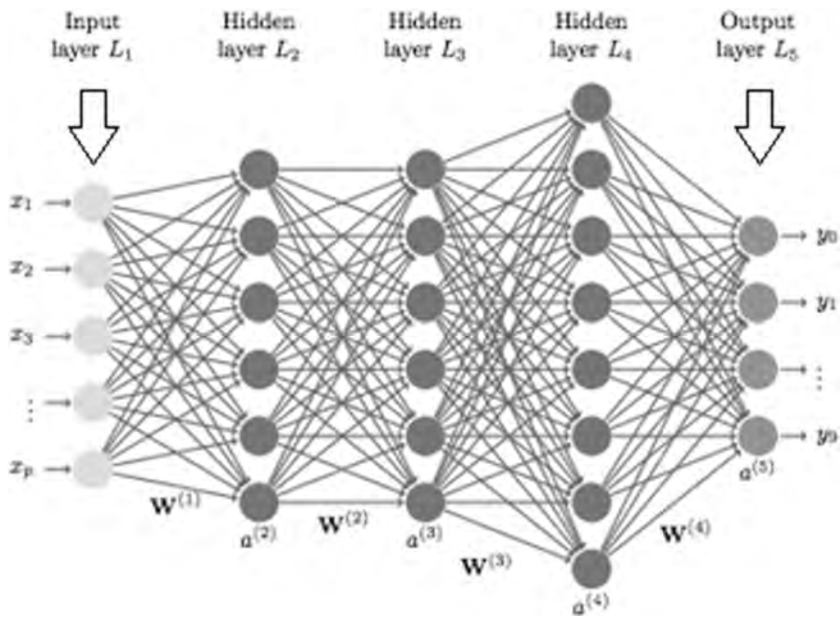


Figure 5.2 An illustration of a deep learning neural network.Source: www.researchgate.net/figure/A-hypothetical-example-of-Multilayer-Perceptron-Network_fig4_303875065.

North Carolina.[1] The speed at which the drones delivered the samples was significantly faster compared to humans, which in this case could mean the difference between life and death. While this use of AI is a positive factor, the initiative instantly eliminated several courier jobs on campus. A few drone developer jobs were created, but the number of jobs eliminated compared to jobs created was far greater. However, expanded technology in any industry can be expected to create similar repercussions. This matter should not be seen as a deterrent from continuing the use of AI in the healthcare industry, as the benefits and capabilities of the technology are just beginning to surface. The possibility of eliminating jobs is always a by-product of innovation and should never be a justification for the status quo. For example, take a look at what Uber has done to the taxicab industry.

Listening Devices and Web Apps

As we investigate the future of AI, consider this scenario:

> While your physician is treating you, an AI listening device in the exam room makes thousands of calculations and predictions

> based on the conversation it picks up between you and the physician. After the exam, the AI device may support the physician's decision-making and/or trigger additional action items, such as ordering the test the physician discussed with you. The AI may also warn the physician of a drug-to-drug interaction, because the AI device already knows what you have purchased over the counter, though you have forgotten to mention it to your physician. Did you know fish oil can be an anti-clotting agent and may negatively affect some treatment plans? Further, the AI already knows you just booked a trip to South America, so it generates a reminder for travel shots. The possibilities are endless.

Google reports that it already has access to 70% of our credit card transactions.[3] Therefore, the AI listening device in the exam room (provided by Google or Amazon) can be linked to the Google database of prior transactions, allowing the real-time conversation with the provider to be reconciled with other key data elements that impact your care. Does this feel a little creepy to you? Are you a bit nervous about data being gathered? Does it make you feel more confident knowing there is a second set of ears, albeit electronic ears, that is helping with decisions? What about all the genetic testing activities that are frequently and widely advertised? What if these data are combined with AI software to achieve a better understanding of past medical history from little-known relatives? The truth is this scenario is already happening all around us.

AI based on crowdsourcing data and/or accessing databases is constantly being utilized all around us. Most everyone has used GPS apps such as Google Maps or Waze. Have you ever seen the apps make real-time suggestions, such as offering new routes based on a set of circumstances like road construction or a traffic jam? Do you ever notice how Facebook and other social media apps seem to know your hobbies, travel preferences and even political affiliations and will make recommendations based on these preferences? Have you seen an ad pop up in your news feed after visiting a new place or restaurant even though you did not disclose your location? Many data collection agencies sell our data, and we often unknowingly release our data on many of the social media and Web-based platforms we regularly visit, such as Facebook, Twitter, Google and Amazon. These companies can see many of the pages you and others are visiting by using hidden tracking technologies, allowing them to tailor their ads accordingly. Another concern that arises is the potential for companies to eavesdrop on phone conversations. While tech companies deny that our phones are listening, devices like Amazon Echo, Alexa and Google Home are becoming increasingly popular. These devices do listen, albeit with the consent of the device's owner. We relinquish a lot of data in exchange for these technology benefits, and we unknowingly consent to giving up these data when we agree to use third-party apps or social medical platforms. Most people never read the lengthy, fine-print user agreements for these platforms, so we unknowingly agree to these terms, which often include stipulations for data collection. Despite a recent increase in public awareness concerning consumer rights and privacy, platform developers stand behind the notion that the use of their applications is a choice we make as consumers. Thus, as a user, we must accept their agreements to access their software.

What happens when these devices start making their way into exam rooms? Your place of work? In a private area? It is one thing for a consumer to invite an *always-listening* artificial intelligence device into their home knowingly, but should we have a right to know when we are being recorded? Under the wire-tap laws, most states prohibit one-party recording, but many companies get around this by having people waive their rights, or they might place a notice of advisement. For example, have you ever heard the statement, "This call may be recorded or monitored for educational purposes?" There is your notice of advisement. Some businesses will post signs letting you know the area is under video surveillance, which may have integrated facial recognition, which opens another world of AI possibilities.

Facial Recognition AI in Healthcare

Another form of AI is facial recognition. Some hospitals are using AI to detect pain or discomfort in patients' facial expressions, or to identify emotions such as stress, depression or anger in a person's voice. This technology is beneficial for nurse lines or suicide crisis centers, but it can also allow an entity to know personal information about each consumer who walks through the door so they can tailor how they interact with you. Some retailers use it to prevent theft by uploading pictures of known shoplifters, and restaurants can identify a big tipper/spender as you enter the premises. A medical practice may use facial recognition to alert the staff about a known hostile patient or for auto check-in, or perhaps to protect sensitive areas such as the maternity ward, allowing access only to known family members.

Most people respond unfavorably when asked how they feel about using facial recognition software that allows other entities to access personal information without consent. However, we have already forfeited a lot of this personal information by agreeing to communicate on the various social media platforms and search engines that collect enormous amounts of data, including images of your face. By now, many of you have uploaded photos to Facebook and noticed how it can automatically tag you in the picture without you making the decision. The same is true with Instagram, Twitter, LinkedIn and others. Other data elements can be associated with these facial images, such as your spending habits, religious affiliations, political views, hobbies, etc. This allows anyone with this software to create customized approaches to how they target you or, worse, treat you differently.

AI in the Exam Room

Today, most healthcare provider organizations have electronic health records (EHR). However, these tools are static databases with algorithms that complement and support humans to complete simple tasks. They do not think for the users; they store data and serve as repositories of information. Now, with the advancement of AI, that factor is changing rapidly. AI today can augment human activity with the ability to sense, understand and learn. AI in healthcare is a collection of technologies that enable machines to understand and learn so they can perform administrative and clinical healthcare functions.

The primary aim of health-related AI applications should be to analyze relationships between prevention and/or treatment techniques and patient outcomes. Privacy policies must catch up to AI to ensure there is no overstepping of boundaries. Thus, the most obvious application of AI in healthcare is data management and its compatibility with the existing EHRs. As with all innovation driven by data, collecting, storing, normalizing and tracing the lineage of the data are the first steps in developing an AI strategy. Today, AI programs are already developed and applied to aid in the diagnosis process, treatment protocol development, drug development, personalized medicine and patient monitoring and care. We are now expecting AI to make its way into eliminating repetitive jobs and to allow for predictive automation.

As AI advances, we will need to consider the creation of ethical standards that are applicable any time patient data are used, with a specific emphasis on patient privacy and accountability for data usage. Now would be a good time to review your patient privacy policies to see how they may need to be updated for changes in your technology, which should also include patient messaging, patient portal, text messaging and others.

The Ethics of AI

Introducing AI into any industry generates several ethical questions:

- What happens when AI shows bias or acts in discriminatory way (see Microsoft's Tay experiment)? An example of this in healthcare could be if the AI started profiling patients with a bad credit score for extra screening before they could schedule an appointment. What if the AI was tracking the race of patients who had balances or missed appointments? AI bias does not have to be intended; it can sometimes evolve based on statistics being collected, especially if the AI is not locked down. For example, let's say 100 patients missed appointments last year. Out of the 100 patients, 30% were born in the month of October. Does this mean that people born in October are awful at keeping commitments, or is this just a random stat with no relevance? The answer would be the latter, but computers don't have critical thinking skills, so it might apply this outcome to future decisions that would treat patients born in October differently.
- How do we counteract inequality and distribute the wealth created by machines? An example of this would be a machine making a clinical diagnosis that would have historically been made by the provider, and the provider would have been paid for his/her expertise to make such a diagnosis. Factors such as the involvement level of the provider and ownership of the AI might be a factor in this example. On the positive side, it could be used to improve capacity, allowing providers to treat patients faster and with better outcomes. As with most innovations that disrupt markets (e.g., Blockbuster vs Netflix), the consumers will respond to whatever creates the greatest value. Capitalizing on it may simply mean embracing it before the competition does. When you consider how disrupters such as Uber and Amazon replaced entire industries overnight, one must always be mindful of change and willing to adapt.
- Can AI have rights? If corporations can be viewed as entities and have rights, can the same be said of machines possessing artificial

intelligence? The legal industry is currently grappling with this question, so this trend is something to watch.

- How do machines affect our behavior and interactions? Many behavioral health experts have already started to express concerns over the impact that computers, especially social media, are having on our culture and behavior. Evidence is mounting about the link between social media and depression. Some feel it distorts reality and has the potential to make one feel inadequate when comparing their own life to images they see on social media, which are often over-hyped. Isolation from human interaction is another concern. Although there are many unknowns, most experts see too much interaction with computers as being unhealthy. Nevertheless, there is an entire market being created to develop companionship bots (also known as sex bots) for those who prefer having a relationship with a computer over a real human.
- Who owns the intellectual property developed by AI once its intended use is achieved? As stated previously, AI is either locked or open AI. A locked AI prevents the AI from adapting, whereas open AI is programmed to evolve and adapt. What happens if/when the AI solves a problem that can later be monetized? How about the data used to power the AI? Should the contributors benefit if it was patient data or consumer input? AI innovation will take input data from many sources. Some argue that platforms such as Facebook, Google and Twitter, who aggregate large quantities of data for financial gains, should share these profits with the consumers/contributors. Others argue that no one is forcing consumers to use or share data on these platforms, so these companies are free to do what they want with the data. The compromise may come down to providing advance notice so consumers can opt in or out. Most consumers are willing to accept the trade-off when they know what to expect.
- In dealing with AI that causes harm, who is at fault? There is currently some movement to mandate that medical AI software and EHRs go through similar FDA approval like most medical devices. The argument for it is to hold vendors accountable for defects that could cause a caregiver to make a mistake. The argument against it is the fact that these are just tools; they don't take the place of judgment. There are also concerns over stifling innovation out of fear of being held accountable for user errors and or an honest defect, which are common when developing new software.

- What happens to privacy and consent when AI is used for tasks like facial recognition? The current HIPAA laws do not address AI facial recognition, although some argue it falls under the same category as recording or video notice requirements. Signage such as, "This area is under video surveillance for your protection," is how these notices are displayed today. Most ignore them or have become insensitive to them.
- Some state governments are looking at enacting the "right to be forgotten" legislation. Under this concept, a consumer can demand the deletion of all their personal data. While this initiative seems logical, can corporations honestly ensure they can purge all the data collected on consumers? What about company records that are necessary for audits, etc.? Let's say a patient demands a provider delete his or her records, but the vendor is in control of the data. Moreover, the provider needs these records for his or her protection. While the "right to be forgotten" is a movement to force technology companies to remove personal data that are stored electronically, the issue is not a black-and-white matter. Thus, the industry needs to improve on allowing patients the rights to correct or amend mistakes in their electronic records.

As industries move forward with implementing AI, these concerns (and others) will continue to be new unknowns and opportunities to adapt for the better.

Moving forward with an AI Strategy

As you begin the journey towards AI, it is critical to keep expectations in perspective. Despite current advances, there are significant challenges in this field that include the challenges and options shown in Table 5.1.

From an adoption standpoint, we recommend the following:

- Learn about and research AI.
- Create awareness among the stakeholders and providers.
- Change potentially negative mindsets toward AI with awareness campaigns that have a focus on benefits and value.
- Educate leadership through attending conferences, workshops, webinars, this publication, etc.
- Set low expectations at the beginning.
- Identify easy wins (most critical).

Table 5.1 AI Challenges and Options for Overcoming

Challenges	*Options for Overcoming the Challenge*
Providers are reluctant to accept AI at the point of care.	Providers will almost always adapt when the benefits and the value for both the patient and caregiver are clear. Focus on the results and benefits. Offer a trial run to demonstrate proof of concept.
Availability of quality data from which to build and maintain AI applications.	At the heart of any AI strategy is the need to have robust data. One option would be to start with the outcome and work back into all the necessary data elements needed to achieve those results. Contrast and compare the current database structure to verify the data are being captured. Most software programs can provide a visual mapping of all the data elements with discrete fields. Vendors rely on subject-matter experts such as providers and nurses to give input on the database structure. For example, what tests typically are ordered for a condition is not something a programmer will know. This must be done as a team consisting of technical and clinical experts.
Missing data streams.	As a follow-up to the above, any gaps in the data elements will need to be addressed. This may mean additional programing or adding the required fields to the intake process.
Limitations of AI methods in health and healthcare software applications.	AI is still in its early stages of development. No one should set unrealistic expectations. Start with something basic leveraging the existing EHR or data repository.

- Embrace cloud computing (see Chapter 2).
- Form an AI workgroup to research and screen vendors.
- Identify gaps in data/capabilities.
- Consider market-based platforms as opposed to self-developed ones.
- Consider enlisting a third-party expert to transfer knowledge and to avoid pitfalls.

Summary

The exponential explosion in the use of AI in healthcare environments compels industry leaders to consider the various applications available for adoption. Although AI delivers many solutions, it is also fraught with multiple problems for patients and providers. Use the ample resources available to educate yourself and your organization on all the potential benefits and pitfalls of the technology. Staying informed and up to date on artificial intelligence will ensure that your organization is using the most current tools available, while adhering to the ethical requirements that this new field imposes.

Chatbot technology today mostly relies on what is called "IF, AND, DO, WHAT LOGIC." If you say THIS, the system will do THAT. An example of this logic might be a patient calling into a nurse line managed by a chatbot. The first interaction with the chatbot would likely include establishing if the person calling is the patient or is calling on behalf of the patient. Next, it might establish if the patient is male or female and if the patient is new or established. If the patient is new, the chatbot interaction will be modified accordingly. Figure 5.3, which shows how patient input and answers are converted to output. The hidden layer is where the programing takes place in accord with the input.

Deep learning uses similar logic as that described above, but there are multiple levels of representation, obtained by composing simple but

Figure 5.3 Chatbot in patient care.

non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. In simple terms, the outcome gets corelated with the desired result. If the result is not achieved, the AI factors this into the programing so that the machine will continue to improve. A well-known (and real-time) example of this can be seen when we use navigation maps on our phones. These navigations maps know that the desired outcome is the fastest route. Therefore the system is always adapting to current situations based on real-time input based on road conditions being sent back by thousands of other drivers traveling these same roads.

AI Resources

The following are several online resources, some free and some with fees, that are available for getting started:

- Udacity's "Intro to Artificial Intelligence" course and the Artificial Intelligence Nanodegree program at www.coursetalk.com/providers/udacity/courses/intro-to-artificial-intelligence.
- Stanford University's online lectures: "Artificial Intelligence: Principles and Techniques" at https://online.stanford.edu/courses/cs221-artificial-intelligence-principles-and-techniques.
- ColumbiaX MicroMasters® Program in Artificial Intelligence online course offered through Columbia University at https://cvn.columbia.edu/content/micromasters-program-artificial-intelligence.
- Microsoft's open-source Cognitive Toolkit (previously known as CNTK) to help developers to master deep-learning algorithms.
- Google's open-source (OS) TensorFlow software library for machine intelligence.
- AI Resources, an open-source code directory from the AI Access Foundation.
- The Association for the Advancement of Artificial Intelligence (AAAI)'s Resources Page.
- MonkeyLearn's Gentle Guide to Machine Learning.
- Stephen Hawking and Elon Musk's Future of Life Institute.
- OpenAI, an open industry and academia-wide deep-learning initiative.

Bibliography

1. Columbus, Louis. 10 Charts That Will Change Your Perspective on Artificial Intelligence's Growth. *Forbes*, January 12, 2019. https://www.forbes.com/sites/louiscolumbus/2018/01/12/10-charts-that-will-change-your-perspective-on-artificial-intelligences-growth/#6f8b0a0e4758. Accessed May 2, 2019.
2. Kellerher, Keven. UPS Begins Using Drones to Transport Medical Samples at North Carolina Hospitals. *Fortune.com*, March 26, 2019. http://fortune.com/2019/03/26/ups-matternet-drones-transport-medical-samples-north-carolina-hospitals/. Accessed May 2, 2019.
3. Reilly, Michael. Google Now Tracks Your Credit Card Purchases and Connects Them to Its Online Profile of You. *MIT's Technology Review*, May 25, 2017. https://www.technologyreview.com/s/607938/google-now-tracks-your-credit-card-purchases-and-connects-them-to-its-online-profile-of-you/. Accessed May 2, 2019.

Chapter Seven

Applied Machine Learning for Healthcare

Prashant Natarajan and Bob Rogers

Computers are useless. They can only give you answers.

—Pablo Picasso

7.1 Introduction

Collecting, managing, and storing big data is a costly exercise if we can't convert such data into high-value, actionable insights or influence workflows in a timely fashion. Generating knowledge from big data increasingly requires the use of machine learning for various reasons—cognitive, organizational, technical, and operational. Any discussion on big data must include a corresponding discussion on machine learning; frankly, they can seldom be separated anymore. "To be useful, data must be analyzed, interpreted, and acted on . . . [and] attention has to shift to new statistical tools from the field of machine learning that will be critical for anyone practicing medicine in the 21st century."[1]

In order to obtain the most value of out of large, diverse, and fast data, we need to consider options beyond rules-based deductive reasoning, "traditional" systems engineering, and descriptive analytics. Artificial intelligence, specifically the sub-fields of machine and deep learning, provides optimal and cost-effective options to expand the universe of knowledge and solutions in healthcare.

---

[1] Obermeyer, Z. and Emanuel, E.J. (2016, September 29). "Predicting the Future—Big Data, Machine Learning, and Clinical Medicine." *New England Journal of Medicine*, Vol. 375, p. 1216.

Machine learning enables new use cases by:

- Ameliorating the effects of certain human limitations—cognitive (repetitive accuracy, human limitations and information overload), physical (fatigue), emotional (mood, human biases, etc.)
- Enabling new knowledge creation or data reduction via learning and prediction
- Learning to generate computational biomarkers—finding hidden patterns/insights that are not visible to the eye
- Processing repetitive data management tasks more efficiently, consistently, and with greater performance
- Serving as the foundation for clinical workflows and comprehensive secondary use that includes predictive and prescriptive analytics, intelligent search, speech-to-text conversion, real-time image processing, among other uses

7.2 Chapter Overview

While there is a plethora of books, videos, websites, and other resources on machine learning, most content is either too rudimentary or, on the other end of the spectrum, requires an advanced understanding of linear algebra, probability, statistics, and/or computer science. In addition, there appears to be a paucity of resources on applied machine learning—in which learning algorithms, data sets, and best practices are optimized for and applied in a specific domain/industry such as healthcare. As with any emerging technology, ensuring the successful design/deployment and "production" use of machine learning requires knowledge that enables you to connect theory to practice and convert general principles into domain-specific applications.

If you are interested in learning more about this exciting field, or just want a better understanding of the truth behind the hype of "how <<Brand X>> <<machine/deep learning>> can cure <<disease 1>>," then this chapter is for you.

7.3 A Brief History

AI and machine learning are not new topics. They have been researched, argued over, and used by computer scientists, applied linguists, engineers, etc. for more than 60 years. The mathematical foundations of machine learning are rooted in algebra, statistics, and probability developed over the last 2000 years. However, modern development of AI and machine learning in the 1950s and '60s began with the works of Alan Turing, John McCarthy, Arthur Samuels, Alan Newell, and Frank Rosenblatt, among others. Samuel's self-learning and optimizing Checkers program is recognized as the first working instance of a machine-learning system. Rosenblatt was instrumental in creating the Perceptron, a learning algorithm inspired by biological neurons that became the basis for the field of artificial neural networks, which we will touch upon later in this chapter. "Feigenbaum and others advocated the case for building expert systems—knowledge repositories tailored for specialized domains such as chemistry and medical diagnosis."[2]

[2] "One Hundred Year Study on Artificial Intelligence, Appendix I, A Short History of AI." Stanford (2016). Available at https://ai100.stanford.edu/2016-report/appendix-i-short-history-ai

In the 1990s, research on machine learning moved from knowledge-engineering–based expert systems to statistical and data-driven approaches. The subsequent time period saw the refinement of backpropagation ("the workhorse algorithm of learning in neural networks"[3]) as also the development of the precursors of what we call *deep learning* today by Hinton and others.[4,5] "Something that can be considered a breakthrough happened in 2006: Hinton et al. [. . .] introduced Deep Belief Networks (DBNs), with a learning algorithm that greedily trains one layer at a time, exploiting an unsupervised learning algorithm for each layer, a Restricted Boltzmann Machine (RBM)."[6]

A more in-depth history of machine learning is beyond the scope of this chapter.

7.4 What's Different About Machine Learning Today?

After many fits and starts over the past decades, machine learning has come out of the hibernation that happened during the "AI winter" that followed the last hype cycle in the 1980s and '90s.[7] Machine learning is also no longer a knowledge-engineering effort as it once was. It's been redefined and optimized to be data intensive instead, hence its appropriateness to handle big data. Today, machine learning (and for that matter, deep learning) is maturing to a point where targeted applications are practical and real. There is definitely increasing market demand, and machine learning is here to stay.

Machine learning is ready for prime time for the following reasons:

1. **Moore's Law.** Continuing advances in computing and storage are allowing us to store and process very large data sets in a cost effective and scalable manner.
2. **Availability of more data.** Machine learning is primarily a data-driven endeavor. As a result, the creation/availability of large data sets coupled with the ability to share/transport such data are allowing us to get further than ever before in predicting or determining new knowledge.
3. **New sources in native unstructured data formats.** Several big-data sources such as the ones discussed in Table 2.1 in Chapter 2 ("Table 2.1 Sources for Big Data in Healthcare" on page 23) are unstructured. Machine learning is ideally suited and is rapidly evolving to better support the generation of insights and analytics directly off native formats such as videos, images, voice, and large un- or semi-structured text.

---

3 Nielson, M. (2016, January). "How the Backpropagation Algorithm Works." Retrieved from http://neuralnetworksanddeeplearning.com/chap2.html

4 Hinton, G.E. and Salakhutdinov, R.R. (2006, July 28). "Reducing the Dimensionality of Data with Neural Networks." *Science,* Vol. 313. Available at https://www.cs.toronto.edu/~hinton/science.pdf

5 Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986, October 9). "Learning Representations by Back-Propagating Errors." *Nature,* Vol. 323, pp. 533–536. Retrieved from http://www.nature.com/nature/journal/v323/n6088/pdf/323533a0.pdf. DOI: 10.1038/323533a0

6 Bengio, Y. (2009). *Foundations and Trends in Machine Learning,* Vol. 2, No. 1, p. 6. DOI: http://dx.doi.org/10.1561/2200000006

7 Katz, Y. (2012) "Noam Chomsky on Where Artificial Intelligence Went Wrong." *The Atlantic* [online]. Available at http://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/?single_page=true

We interact with machine learning (and learning algorithms) on a daily basis; examples include self-driving cars, email spam filters, Netflix movie suggestions, Amazon shopping recommendations, and postal-code–based mail sorting using handwriting recognition. Machine learning applications are rapidly being deployed and used in the commercial space across diverse verticals—retail and e-commerce, government, finance, healthcare (providers, payers, and pharma, and personal/public health), cyber security, transportation, agriculture, space exploration, and manufacturing, among many others.

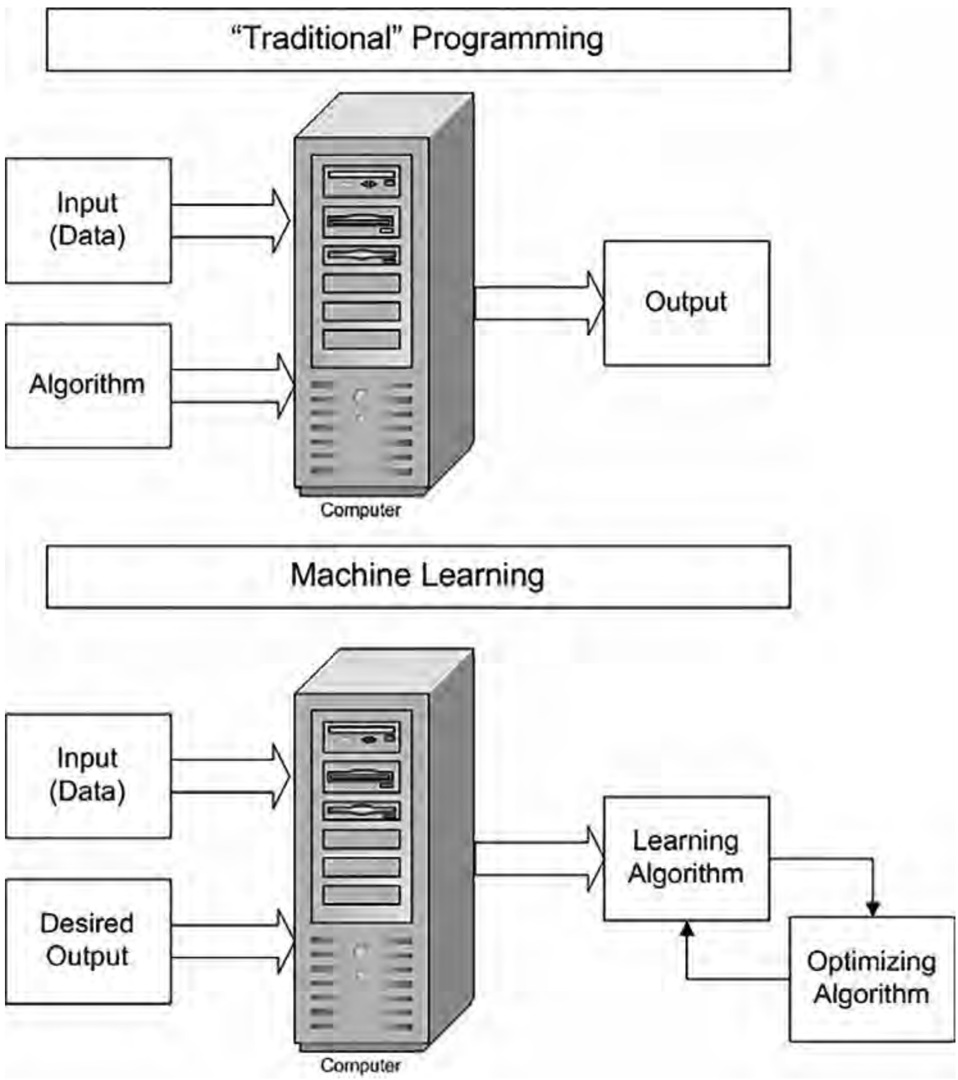


Figure 7.1 Traditional programming and machine learning: a comparison.

7.4.1 What Is Machine Learning?

Arthur Samuel is credited with defining machine learning as the the field of study that gives computers the ability to learn without being explicitly programmed. While its simplicity ensures that this definition is oft-quoted, others provide added perspectives and useful clarifications:

> Machine Learning is a paradigm that enables systems to automatically improve their performance at a task by observing relevant data. Indeed, machine learning has been the key contributor to the AI surge in the past few decades, ranging from search and product recommendation engines, to systems for speech recognition, fraud detection, image understanding, and countless other tasks that once relied on human skill and judgment.[8]

> One useful perspective on machine learning is that it involves searching a very large space of possible hypotheses to determine one that best fits the observed data and any prior knowledge held by the learner.[9]

As we will review in the next section, machine learning is different from traditional software programming due to its emphasis on:

- Learning algorithms versus "traditional" algorithms
- Reasoning that is primarily induction and abduction, with a selective emphasis on deduction
- Dealing with uncertainty ("the unknown unknowns") via the use of mathematical models that are driven by probability and statistics as compared to deterministic rules
- Prediction: using data you have to extrapolate data you don't have in order to infer probability of outcomes

7.5 How Do Machines Reason and Learn: A Crash Course in Learning Algorithms

"A learning algorithm is an algorithm that is able to learn from data."[10] Machine learning is different from traditional programming (see Figure 7.1). "In machine learning, we provide the input (data), the desired result and out comes the [learning] algorithm."[11] Learning algorithms—also known as *learners*—are algorithms that create new knowledge or demonstrate new skills by learning from old (training) data and new (generalized) data. A learning algorithm uses data and experience to self-learn and also to perform better over time. During the process, a learner also optimizes itself to progressively come up with better predictions (see Figure 7.1).

8 "One Hundred Year Study on Artificial Intelligence, Appendix I, A Short History of AI." *op. cit.*
9 Mitchell, T.M. (1997). *Machine Learning.* McGraw Hill. p. 27.
10 Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning.* Book in preparation for MIT Press. Information available at http://www.deeplearningbook.org. p. 98.
11 Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World.* Basic Books. p. 5.

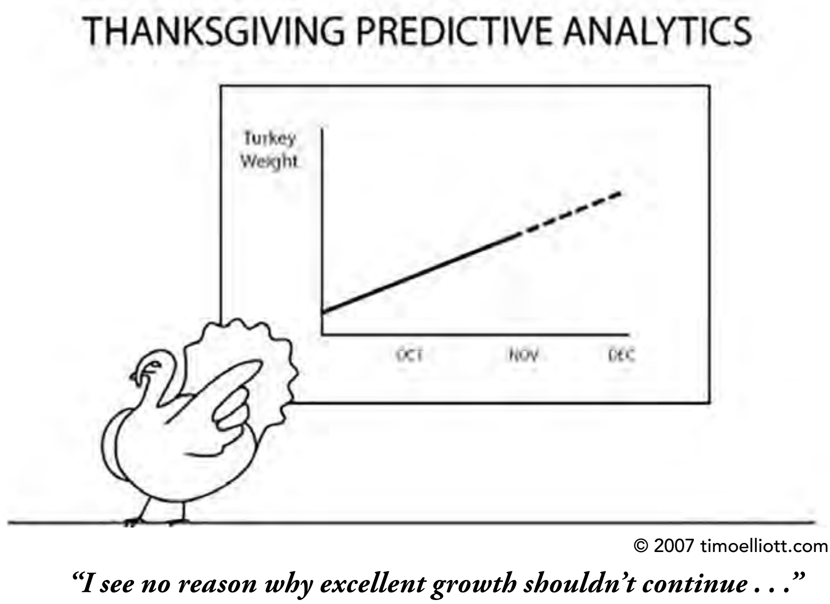


Figure 7.2 Prediction is not certainty. (*Source:* Timo Elliot; used with permission.)

Learners are the foundation of any machine-learning system, and they help us achieve generalization via induction or abduction. *Generalization* is a core concept in machine learning; to be useful, machine-learning algorithms can't just *memorize* the past, they must *learn from* the past. Generalization is the ability to respond properly to new situations based on experience from past situations.

We will introduce two basic concepts here: *Training Dataset* (the data you have that is used as input to the learner to train the model), and *Test Dataset* (dataset is used by the learner for validation and optimization).

Training model refers to the ML artifact that is created coming out of the training process. Training is supplying the learning algorithm with training data to learn from.

A *cost function* is something (usually, a function) that you want to minimize in the ML system. For example, your cost function might be the sum of squared errors over your training data set.[12]

In summary, a machine-learning system consists of the following basic components:

- Learning and optimizing algorithms
- Training and test datasets
- Training model
- Cost function[13]

[12] "Can someone explain to me the difference between a cost function and the gradient descent equation in logistic regression?" (2012). StackOverflow, StackExchange. http://stackoverflow.com/questions/13623113/can-someone-explain-to-me-the-difference-between-a-cost-function-and-the-gradien

[13] Goodfellow, I., et al. *op. cit.* p. 99.

Table 7.1 Some Example Learning Problems[a]

Example Learning Problem	Task T	Performance Measure P	Training Experience E
Learning Checkers	Playing checkers	% of games won against opponents	Learner playing practice games against itself
Handwriting recognition	Recognizing and classifying handwritten words within images	% of words correctly classified	Training dataset of handwritten words with given classifications
Self driving car	Driving from Cupertino to Livermore on public roads	Average distance travelled before an error (as judged by a human overseer)	Sequence of videos, still images, and steering commands recorded while observing a human driver

[a] Table adapted from Mitchell, T.M. (1997). *Machine Learning.* McGraw Hill. pp. 3–4.

In summary, "a well-defined learning problem requires a well-specified task, T; performance metric, P; and source of training experience, E."[14] A formal (and personal favorite) definition of learning states, "A computer program 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."[15]

7.6 Mastering the Basics of Machine Learning

Now that we've reviewed Mitchell's definition, let's take a look at Task T, Performance P, and Experience E using examples in Table 7.1.

7.6.1 Task, T

A *Task* is something that we want the machine learning system to do: "the process of learning itself is not the task. Learning is our means of attaining the ability to perform the task."[16] To illustrate further, a task for a self-driven car would be to make the journey autonomously from Peoria, Illinois, to Livermore, California. The learning algorithms and the rest of the machine learning system that the car uses to learn and recognize street signs, sidewalks, other vehicles, people, etc. constitute the means by which this task is completed successfully. Examples of tasks that can be done by a machine-learning system are classification, classification with missing inputs, regression, transcription, machine translation, structured output, anomaly

[14] Mitchell, T.M. *op. cit.*
[15] *Ibid.* p. 2.
[16] Goodfellow, *op. cit.* p. 99.

detection, synthesis, density estimation, denoising, imputation of missing values, named entity recognition, etc.[17]

In machine learning, "a *Dataset* is a collection of Examples."[18] "A *Task* is defined in terms of how the learning system should process an Example. An *Example* is defined as a collection of features that have been quantitatively measured from some object or event that we want the system to process. A *Feature* is the combination of an attribute and its value." For instance, "Color is blue" is a feature, where color is the attribute and blue is the value. Another example: "The features of an image are usually the values of the pixels in the image."[19]

Note that some authors prefer to use the terms *Feature* and *Example* as synonyms.

In machine learning, we are very interested in understanding and addressing *dimensionality,* which is defined as the number of features that contain most useful or actionable information, as well as *Parameters,* which are attribute values (of high-value features) that control the behavior of the learning system. Parameters are important, as they can be modified by the learning algorithm to determine better performance (accuracy) or contextualization of prediction. So for example, the predicted sale price of a house may be impacted more by parameters such as nearness to school or number of bedrooms, rather than the type of roof shingles or color of exterior paint.

As we saw in earlier chapters, with big data, it has become easier to collect/store/manage data than to obsess over or minimize what is collected. "We are in the era of massive automatic data collection, systematically obtaining many measurements, not knowing which ones will be relevant to the phenomenon of interest. Our task is to find a needle in a haystack, teasing the relevant information out of a vast pile of glut. This is a big break from the past, when it was assumed that one was dealing with a few well-chosen variables—for example, using scientific knowledge to measure just the right variables in advance."[20]

Understanding Dimensionality allows us to discuss *dimensional reduction,* which is reducing the number of features required in an example. Dimensional reduction refers to the algorithmic processes by which a dataset having high dimensions is converted into a dataset with fewer dimensions so that as much information as possible about the original data is preserved. This is done in machine learning to address what is known as Bellman's *Curse of Dimensionality,* in which "many learning algorithms that work fine when the dimensions are low become intractable when the input is high-dimensional. But in machine learning [the curse] refers to much more. Generalizing correctly becomes exponentially harder as the dimensionality (number of features) of the examples grows."[21]

> Dimension reduction can be accomplished via *Feature Extraction* and *Feature Selection.* Feature extraction creates new features resulting from the combination of the original

[17] *Ibid.* pp. 100–103.

[18] *Ibid.* pp. 104.

[19] *Ibid.* pp. 99.

[20] Donoho, D.L. (2000). "High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality." Stanford University. Retrieved from http://statweb.stanford.edu/~donoho/Lectures/CBMS/Curses.pdf. p. 17.

[21] Domingos, P. (2012, October). "A Few Useful Things to Know about Machine Learning." *Communications of the ACM,* Vol. 55, No. 10, p. 81.

> features; and feature selection produces a subset of the original features. Both attempt to reduce the dimensionality of a dataset in order to facilitate efficient data processing tasks.[22]
>
> *Weights* determine how each feature affects the prediction. If a feature receives a positive weight, then increasing the value of that feature increases the value of our prediction. If a feature receives a negative weight, increasing the value of that feature reduces the value of our prediction. If the weight is 0, there is no effect on prediction.[23]

One of the primary purposes of machine learning is for the system to perform on "unknown unknowns," or new, previously unseen data—not just the training dataset with which the model was trained. A good machine-learning system generalizes well from the training dataset to any data from the problem domain. This allows the system to extrapolate and predict on "never seen before" data. Wilson explains further:

> Learning in backprop seems to operate by first of all getting a rough set of weights which fit the training patterns in a general sort of way, and then working progressively towards a set of weights that fit the training patterns exactly. If learning goes too far down this path, one may reach a set of weights that fits the idiosyncrasies of the particular set of patterns very well, but does not interpolate (i.e., generalize) well.[24]
>
> Generalization is the ability to perform well on previously unobserved inputs. Generalization Error (also called Test Error) is defined as the expected value of the error on a new input.[25]

Overfitting has significant impacts on the performance of the machine learning system. Overfitting happens when a learner mimics random fluctuations, anomalies, and noise in the training dataset, thus adversely impacting the performance of the system on new data. "[W]ith large complex sets of training patterns, it is likely that some errors may occur, either in the inputs or in the outputs. In that case, it is likely that [the learner] will be contorting the weights so as to fit precisely around training patterns that are actually erroneous."[26]

7.6.2 Performance, P

In order to evaluate the abilities of a machine-learning algorithm, we must design a quantitative measure of its *Performance, P.* Performance is usually measured on the task being carried out by the machine-learning system and is typically measured in terms of *Accuracy,* which is the "proportion of examples for which the model produces the correct output," or *Error Rate,* which is the "proportion of examples for which the model produces the incorrect output."[27]

---

[22] Dash, M. and Liu, H. (n.d.). "Dimensional Reduction." Retrieved from: http://www.public.asu.edu/~huanliu/papers/dm07.pdf

[23] Goodfellow, I., et al. *op. cit.* pp. 107–108.

[24] Wilson, H.B. (1998, updated June 24, 2012). *The Machine Learning Dictionary.* http://www.cse.unsw.edu.au/~billw/mldict.html#generalizebp

[25] Goodfellow, I., et al. *op. cit.* p. 110.

[26] Wilson, B. *op. cit.*

[27] Goodfellow, I., et al. *op. cit.* pp. 103–104.

We conclude our brief discussion on performance by reviewing *Noise,* which in machine learning refers to "errors in the training data for machine learning algorithms. If a problem is difficult enough and complicated enough to be worth doing with machine learning techniques, then any reasonable training set is going to be large enough that there are likely to be errors in it. This will of course cause problems for the learning algorithm."[28]

7.6.3 Experience, E

Experience in machine learning is primarily determined by the amount of supervision (during the learning process) and the availability of labeled data in the dataset.

In *supervised learning,* "algorithms experience a dataset containing features but each example is associated with a *Label* (AKA *Target*)."[29] Wilson defines supervised learning as "a kind of machine learning where the learning algorithm is provided with a set of inputs for the algorithm along with the corresponding correct outputs, and learning involves the algorithm comparing its current actual output with the correct or target outputs, so that it knows what its error is, and modify [*sic*] things accordingly."[30] Input data is labeled based on existing knowledge (for example, is the email in the training dataset spam or not-spam?) The model continues to train until it achieves a desired level of performance on the training dataset, and the training model is then fed new and unknown data, as described earlier.

In *unsupervised learning,* input data is not labeled, and furthermore, "the system is not told the 'right answer'—for example, it is not trained on pairs consisting of an input and the desired output. Instead the system is given the input patterns and is left to find interesting patterns, regularities, or clusterings among them."[31]

In *semi-supervised learning,* as the experience suggests, input data may be only partially labeled, and the expected results may or may not be known. The machine learning system will include both supervised and unsupervised learners.

Active learning is a semi-supervised learning experience in which "the model chooses by itself what unlabelled data would be most informative for it, and asks an external 'oracle' (for example, a human annotator) for a label for the new data points."[32] The learner aims to "achieve high accuracy using as few labeled instances as possible, thereby minimizing the cost of obtaining labeled data (something that remains challenging in healthcare)."[33]

Deep learning is a type of machine-learning experience that uses learning algorithms called *artificial neural networks* that attempt to simulate or replicate the functioning of the human

---

[28] Wilson, B. *op. cit.* http://www.cse.unsw.edu.au/~billw/mldict.html#firstN

[29] Goodfellow, I., et al. *op. cit.* pp. 105.

[30] Wilson, B. *op. cit.* http://www.cse.unsw.edu.au/~billw/mldict.html#firstS

[31] *Ibid.*

[32] Gal, Y. (2016). *Uncertainty in Deep Learning,* PhD Thesis, University of Cambridge. Retrieved from http://mlg.eng.cam.ac.uk/yarin/blog_2248.html. p. 11.

[33] Settles, B. (updated 2010, January 6). "Active Learning Literature Survey." *Computer Sciences Technical Report 1648,* University of Wisconsin–Madison. Retrieved from http://burrsettles.com/pub/settles.activelearning.pdf

brain. Think of deep neural networks as "ANNs with lotsa depth."[34] Before we review deep learning, let's take a quick look at artificial neural networks and understand why they serve as a basis for understanding what deep learning does.

7.7 Artificial Neural Networks: An Overview

- **Biological neuron.** "From the artificial neural network point of view, a biological neuron operates as follows: electrical pulses from other neurons cause the transfer of substances called neurotransmitters (of which there are several varieties) from the synaptic terminals of a neuron's axon (think "output") across a structure called a synapse to the dendrites of other neurons (call them downstream neurons). The arrival of the neurotransmitter in the dendrite of the downstream neuron increases the tendency of the downstream neuron to send an electrical pulse itself ("fire"). If enough dendrites of a neuron receive neurotransmitters in a short enough period of time, the neuron will fire."[35]
- **Artificial neuron.** "A simple model of a biological neuron used in neural networks to perform a small part of some overall computational problem. It has inputs from other neurons, with each of which is associated a weight—that is, a number which indicates the degree of importance which this neuron attaches to that input."[36]
- **Artificial neural network.** "An artificial neural network is a collection of simple artificial neurons connected by directed weighted connections. When the system is set running, the activation levels of the input units is clamped to desired values. After this the activation is propagated, at each time step, along the directed weighted connections to other units. The activations of non-input neurons are computing using each neuron's activation function. The system might either settle into a stable state after a number of time steps, or in the case of a feed forward network, the activation might flow through to output units. Learning might or might not occur, depending on the type of neural network and the mode of operation of the network."[37]

7.8 Deep Learning

"*Deep learning* is a specific kind of machine learning. In order to understand deep learning well, one must have a solid understanding of the basic principles of machine learning."[38] It is a "kind of learning where the representations you form have several levels of abstraction, rather than a direct input to output."[39] Think of "deep" in deep learning as having many more layers

---

[34] @natarpr (author) on Twitter. (2016, December 8).
[35] Wilson, B. *op.cit.* http://www.cse.unsw.edu.au/~billw/mldict.html#bioneuron
[36] Wilson, B. *op.cit.* http://www.cse.unsw.edu.au/~billw/mldict.html#neuron
[37] *Ibid.*
[38] Goodfellow, I., et al. *op. cit.* p. 98.
[39] Norvig, P. (2016, March 18). "Deep Learning and Understandability versus Software Engineering and Verification." Available at http://youtu.be/X769cyzBNVw

(or *Depth*) than were possible with ANNs and as the ability to deal with very large datasets due to Moore's law and data availability. The principle driving deep learning is "guiding the training of intermediate levels of representation using unsupervised learning, which can be performed locally at each level."[40]

Deep learning particularly does well on sequential, unstructured, or analog data such as images, audio, and video and is becoming very popular today due to its high performance. "Deep Learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer."[41]

Deep learning currently excels at supervised learning. However, we see as much or greater potential for using deep learning in unsupervised learning, "primarily because large datasets contain greater amounts of unlabeled data that require labeling, which is time- and effort-intensive."[42]

Let's discuss some types of deep neural nets, including Feed Forward Neural Networks; Recurrent Neural Networks; Convolutional Neural Networks; and Reinforcement Neural Networks.

- **Feed Forward Neural Network.** A "kind of neural network in which the nodes can be numbered, in such a way that each node has weighted connections only to nodes with higher numbers. [. . .] In practice, the nodes of most feedforward nets are partitioned into layers—that is, sets of nodes, and the layers may be numbered in such a way that the nodes in each layer are connected only to nodes in the next layer. The first layer has no input connections and is termed the input layer. The last layer has no output connections and is termed the output layer. The layers in between the input and output layers are termed hidden layers, and consist of hidden units."[43]
- **Recurrent Neural Network.** Sequence-based neural networks that play a key role in natural language processing, machine translation, video processing, and many other tasks.[44] "The idea behind RNNs is to make use of sequential information. In a traditional neural network, we assume that all inputs (and outputs) are independent of each other. But for many tasks that's a very bad idea. If you want to predict the next word in a sentence you better know which words came before it. RNNs are called *recurrent* because they perform the same task for every element of a sequence, with the output being depende[nt] on the previous computations. RNNs [. . .] have a 'memory' which captures information about what has been calculated so far."[45]

---

40 Bengio, Y. (2009). "Learning Deep Architectures for AI." *Foundations and Trends® in Machine Learning,* Vol. 2, No. 1, p. 7. Information available at http://dx.doi.org/10.1561/2200000006

41 LeCun, Y., Bengio, Y., and Hinton, G. (2015 May 28). "Deep Learning." *Nature,* Vol. 521, pp. 436–444. Retrieved from http://www.nature.com/nature/journal/v521/n7553/abs/nature14539.html

42 Ng, A. (2014, August 26). "Deep Learning: Machine Learning via Large-scale Brain Simulations." Invited Talk: Deep Learning, Stanford University. http://youtu.be/W15K9PegQt0

43 Wilson. *op. cit.* http://www.cse.unsw.edu.au/~billw/mldict.html#firstF

44 Yarin, G. *op. cit.* p. 5.

45 "Recurrent Neural Networks Tutorial, Part 1: Introduction to RNNs." (2015, September 7). Retrieved from http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/

- **Convolutional Neural Network.** A "type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex."[46] CNNs excel at dealing with sequential, analog, or unstructured data and are showing great promise in healthcare—particularly in image/audio/video recognition, recommender systems, and natural language processing. The model is made of a "recursive application of convolution and pooling layers, followed by simple NNs. A convolution layer is a linear transformation that preserves spatial information in the input image. Pooling layers simply take the output of a convolution layer and reduce its dimensionality."[47] An excellent example of how deep learning and CNNs are being used in healthcare is the work being done by Pratik Mukherjee MD and his team at UCSF (see UCSF Case Study on page 149).

- **Reinforcement Neural Network**. Neural networks in which the learner learns and performs tasks via trial and error, much like a child learning to ride her bicycle. Reinforcement learning is inspired by behaviorist psychology and focuses on how software agents ought to take *actions* in an *environment* so as to maximize some notion of cumulative *reward.*

  "Reinforcement learning differs from supervised learning in that correct input/output pairs are never presented, nor sub-optimal actions explicitly corrected."[48]

7.9 A Guided Tour of Machine-Learning Algorithms in Healthcare

Every machine-learning algorithm is good at answering a specific kind of question. Let's take a look at some of the most important algorithms in healthcare and the questions they are being used to answer.

Don't be intimidated by the sheer number of machine-learning algorithms that are out there. While there are currently 54 Wikipedia pages dedicated to specific machine-learning algorithms, all of these are really variations on a few major themes.

The list of algorithms described below is not exhaustive. Our goal is to give examples of the most commonly used methods to answer different types of questions in healthcare analytics. In addition to the type of question being answered, there is another useful way to characterize machine-learning algorithms: whether or not they need input data that is labeled with known answers to create them. Methods that require input data with known labels are called *supervised training* algorithms, and those that do not require any prior knowledge of what answers are expected are called *unsupervised.* The majority of the algorithms below are supervised learning algorithms. We indicate the unsupervised learning algorithms with an asterisk (*).

7.9.1 Classifier

Does data belong to class A?
Example: Is this really a heart failure patient?

[46] "Convolutional Neural Network." https://en.m.wikipedia.org/wiki/Convolutional_neural_network
[47] Yarin, G. *op. cit.*
[48] "Reinforcement Learning." (n.d.) Wikipedia. https://en.m.wikipedia.org/wiki/Reinforcement_learning

- **Logistic regression.** Logistic regression is the workhorse of classifiers. It is a linear classifier, which means it uses a single, straight cut to divide the world of possible features into two groups. If a patient's characteristics fall on one side of this cut, they are in class A (i.e., they are judged to have heart failure), otherwise they are not in class A. In problems with many features (some problems can require millions of features), logistic regression is the preferred method because it works well and is straightforward to compute.
- **Support Vector Machine (SVM).** SVM is a linear classifier with a twist: the world of possible features is split by a single line as in logistic regression, but this line can be curved. This additional flexibility makes SVM highly adaptable, but because of the way the curvature is introduced (though something called a kernel), they are still simple to compute and to interpret.
- **Decision tree, random forest, boosted trees.** Trees and forests are an entire family of algorithms, all based on the idea of creating a tree of decisions about features that lead to a specific classification. For example, to identify heart failure, the algorithm may start with ejection fraction. Is ejection fraction above or below 50? For each of these paths, a new question would be considered, such as: does the echocardiogram show Left Ventricle Hypertrophy? At the end of each series of questions, the patient falls into either a "heart failure" or a "not heart failure" bucket. Random forests improve upon decision trees by dividing the input data into many different subsets and creating a different decision tree for each of these subsets. All of the different resulting decision trees then vote to determine the final classification of the input. This process reduces the risk of making the final buckets too small and subsequently being fooled by random variations in the original labeled training data. Boosting is a trick for creating decision trees and random forests that can significantly improve their ability to generalize from example data. We call them out specifically here because boosted tree classifiers tend to be among the best-performing algorithms in public classification competitions such as Kaggle (www.kaggle.com).
- **Deep Learning.** Deep learning, and indeed neural networks in general, can take raw data as input and produce a class (or a vector of probabilities for many classes) as output. All neural network models consist of multiple layers of "neurons": each neuron in a layer receives the outputs of neurons in previous layers, combines these inputs, and uses a threshold to determine whether to output a value closer to 0 or closer to 1 for processing by the next layer. Deep-learning algorithms are unique in their ability to automatically generate features of interest in input data, as long as they are provided with a sufficient number of training examples (usually in the millions). Deep learning is already extensively used in image, video, and audio understanding in healthcare, and it will eventually become more common in other classification problems in healthcare as larger sets of labeled training data become available for healthcare applications.

Common uses: Classifiers are the most commonly used machine learning algorithms in all analytics applications, including healthcare. In healthcare, classifiers are used to:

- Suggest possible patient diagnoses
- Identify patients with high readmission risk

- Automatically alert care providers early in the development of sepsis
- Define the thresholds for "abnormal" lab results
- Automatically differentiate between clinical and administrative documents
- Recommend the most effective wellness or disease management intervention for a patient
- Many, many more

7.9.2 Memory-Based Learning*

How does this new piece of data compare to past data?
Example: Who are the patients most like this patient?

- **Associative memory.** An associative memory system compares incoming data with past data to identify what the new data is most like. The comparison can be based on any subset of the attributes of the data, so no assumptions need to be made about what is "important" in the data, and very large numbers of features can be included. This makes these algorithms especially useful in healthcare applications, because the number of conditions and measurement results that could be applicable to a patient is very large.

Common uses: Memory-based learning is commonly used in healthcare to:

- Create cohorts of patients with which to compare a specific patient. This is the "patients like mine" question that plays a role in treatment planning, pharmaceutical research, and risk adjustment modeling.
- Identify insurance fraud.
- Calculate risk of readmission and other costly future events.

7.9.3 Topic Modeling

What is this document about?
Example: What conditions are being addressed for this patient in this SOAP note?

- **Latent Dirichlet Allocation (LDA).** LDA assumes that content is made up of a combination of underlying topics. A single doctor's note may be 80% about a patient's diabetes and 20% about pain management. LDA can identify the combinations of terms and phrases that make up the underlying topics. LDA can be applied to many different sources of information, from single documents to groups of documents, to even a patient's entire clinical history.
- **Probabilistic Latent Semantic Analysis (pLSA), Latent Semantic Analysis (LSA).** These algorithms are similar to LDA but make stricter simplifying assumptions about how topics are distributed in documents and how words are distributed in topics. With modern computing and large datasets available for analysis, these assumptions are no longer necessary, so LDA is the dominant methodology.

Common uses in healthcare:

- Reliably identify the conditions that a patient has based on clinical text combined with structured data for use in acute disease detection, such as:
  - Sepsis detection
  - Heart failure detection for prevention of hospital readmission
  - Drug-seeking and drug fraud

7.9.4 Forecasting

How much will this time series change in the next time period?
Example: How likely is this CKD patient to progress in the next six months?

- **Linear regression.*** Draws a straight line through the time series of past data, assuming that the current linear trend will continue. This approach requires that the predicted output is a continuous variable.
- **Neural networks.*** A neural network, which can be as simple as a Multi-Layer Perceptron (MLP) or as complex as a recurrent deep-learning model (e.g., Long Short-Term Memory, LSTM), takes past values as inputs and produces the predicted next value as output. All neural network models consist of multiple layers of "neurons": each neuron in a layer receives the outputs of neurons in previous layers, combines these inputs, and uses a threshold to determine whether to output a value closer to 0 or closer to 1 for processing by the next layer.
- **Exponential smoothing.*** This is a simple but surprisingly useful method for predicting the next value in a time series based on a weighted average of the most recent past values. It gives the most weight to the most recent measured value, then reduces the weight by multiplying by a number between 0 and 1 for each subsequent previous value, resulting in an exponential decrease in the impacts of older previous time-series values.
- **Auto-Regressive Integrated Moving Average (ARIMA) modeling.*** This is a general group of forecasting methods (of which exponential smoothing is actually a member) that uses weighted averages of past time-series values, past differences between time-series values, past differences between rates of change of past values, and so on, to calculate future values of the time series.

Common uses in healthcare: The expected next value of a time series is often used as part of a larger predictive modeling or clinical decision-support application. For example:

- Chase lists for disease management: the predicted future values of key diagnostic measurements such as Hemoglobin A1C or creatinine are used to determine who should be included on a chase list for chronic disease management.
- Risk prediction for individuals: a number of healthcare companies, payors, providers, and third-party analytics vendors use predictive models to compute the likelihood that a patient will convert to a new diagnosis within future time periods ranging between six months and two years.

- Very time-sensitive detection of disease: acute applications, such as sepsis detection in the hospital, commonly include time-series forecasts of key measurements such as temperature, white blood cell count, or respiratory rate.

7.9.5 Probability Estimation

What is the most likely interpretation of the data?
Example: What is the most likely diagnosis, given the patient's signs, symptoms, and measurements?

- **Probabilistic Graph Model (PGM).** PGM algorithms, such as Bayes networks, identify key observations, measurements, and outcomes and link them together to identify causal relationships. Each of these factors would be represented as a node in a graph, with connections between nodes indicating causal relationships. These graphs can be learned directly from data or constructed by human experts. The PGM algorithm then uses actual data to determine the amount of influence each combination of variables (nodes) has on the others.
- **Logistic regression.** Logistic regression models assume that the log of the odds of an event occurring (or an interpretation being applicable) can be calculated from a simple weighted average of a set of observations. In practice, this means that they can be used to predict the likelihoods of categorical values such as diagnoses or specific outcomes.

Common uses in healthcare: Sepsis detection, readmission prevention.

- CDS and diagnosis tools: For example, a model for diagnosing COPD might include historical and demographic information such as age, sex, smoking history, exposure to chemicals, and signs and symptoms such as coughing, dyspnea, and blood oxygen saturation, along with comorbidities such as bronchitis, diabetes, and lung cancer. Each of these factors would be represented as a node in the graph, with connections between nodes indicating a causal relationship. The presence or absence of a combination of these factors will influence the probability that a COPD diagnosis is applicable to the patient.
- Disease risk forecasting: PGM and logistic regression can both be used to compute the probabilities of different diagnoses or interpretations of data.

7.9.6 Image and Video Understanding

What is in this image? What is happening in this video?
Example: Is there a nodule in this chest x-ray?

- **Deep learning.** Deep-learning systems, especially convolutional neural networks (CNNs), are very powerful methods for recognizing objects or patterns in complex images. The power of deep-learning algorithms is that, given enough data, they can learn what is important for understanding an image without being explicitly told. In practice, to recognize a cat in a photo, or a nodule in a chest x-ray, a deep-learning system may need to be shown millions of images, each labeled with the desired answer for that image. This training process can be very computationally intensive and require long times to complete (hours, days, and even weeks), but once the algorithm is trained, it can easily be used to quickly recognize the objects it has been taught—a process called "inference" or "scoring."

Common uses in healthcare: This area is exploding right now. At the time of writing, there are compelling deep-learning results being developed for:

- Automated detection of "findings" in radiology images: for example, features such as nodules, pneumonia, or pneumothorax can be automatically detected in chest x-rays using deep-learning systems. These results can be used to route time-sensitive cases to radiologists or to enhance the productivity of radiologists without sacrificing accuracy. Development is underway to develop commercially viable radiology detection systems for all modalities, including MRI, CT, and ultrasound.
- Workflow monitoring and procedural compliance: prototype systems have been developed that can use video and other data streams (such as RFID-based location tracking) to track compliance with standard workflows and procedures. For example, nosocomial infection prevention, in which deep-learning systems have been developed to recognize activities in hospitals that increase the risk of nosocomial infection. These systems, for instance, can flag when a wound is handled but hands are not washed before an IV is placed.
- Patient safety monitoring: Patients can be monitored via video to predict their fall risk in general and to identify when they are getting out of the bed or performing a risky activity, so that personnel can be alerted to assist.

7.9.7 Speech to Text

What is the transcribed text for this audio stream?
Example: What did the clinician dictate?

- **Hidden Markov Models (HMMs).** HMMs assume that there are underlying processes that we can't see—but which are nonetheless consistent and predictable—that create outputs that we can see. For example, in a sentence, if the word "mellitus" is detected, the previous word is far more likely to be "diabetes" than it is to be "disabilities." This is very valuable in speech-to-text processing, because it is not possible to clearly hear or identify each word in an audio stream. The HMM can help choose the right interpretation of the sounds to result in the correct overall transcription. HMM has been used historically in a number of commercial dictation transcription systems.
- **Deep learning, especially Long Short-Term Memory (LSTM).** LSTM models, like the Convolutional Neural Networks described above, can automatically learn what attributes of an audio stream are important for predicting what words it represents. Given sufficient data, which is readily available to online service providers such as Google and Baidu, it is possible to train LSTM models to accurately convert spoken language into

text in almost any language. This technology has become the state of the art for spoken language understanding applications and will likely play an increasing role in clinical transcription applications.

Common uses in healthcare:

- Dictation and clinical note transcription.
- Interpretation and automated documentation of clinical encounters, including speech from clinicians, patients, and support staff.
- Voice controls for computer systems in the clinic and in the surgical theater.
- Call-center resources and agent coaching to help call-center operators provide appropriate information and resources to patients or members.
- Patient coaching: applications are being developed in which patients are given context-dependent coaching for disease management, wellness, and particularly behavioral health applications.

7.9.8 Recommender Systems

What was the behavior of other people like you?
Example: What chronic disease management intervention is most likely to be effective for this patient?

- **Collaborative filtering.** Collaborative filtering includes several different methods for predicting a user's rating for a specific item given the user's history of ratings for other items, combined with the history of all users' ratings for all items. Intuitively, if a user rates an item highly, then that user is likely to give a high rating to a very similar item. For healthcare, users could be patients, items could be treatments or interventions, and ratings could be outcome or level of compliance.
- **Memory-based learning.** Memory-based learning systems, such as Saffron, compute the difference between a new data point and previously seen data, for a number of different contexts. When the new data is near previous data, it is possible to predict the outcome based on what happened in the past. These systems tend to learn continuously on an ongoing basis as they are exposed to more data, and they can be used to reason on very complex data.
- **Association rules.** Association rules are a data-mining method in which algorithms use historical data to identify items or events that commonly occur together. For example, a patient needing health education and weight management is also highly likely to need nutrition management.

Common uses in healthcare:

- Matching patients with interventions and coaching resources
- Detecting fraudulent claims
- Call-center optimization and customer experience

7.9.9 Clustering

Can the data be grouped into natural categories or buckets?
Example: Are there natural groupings that can help me understand my patients?

- **Unsupervised Clustering.*** Unsupervised clustering algorithms, such as K-Means, can automatically identify naturally occurring groups of similar items. Typically, the algorithm is given a set of attributes for each item (for example, diagnoses and lab measurements for each patient) and a number of clusters to create. The algorithm will then work out which combinations of attributes most accurately divide the items into that number of groups. The resulting "clusters" can usually be interpreted by humans by looking at which attributes are most important in the cluster.
- **Hierarchical clustering.*** This family of methods creates a tree or dendogram of clustering scenarios for data, creating a single cluster containing all the items, which then splits into two clusters, each of which further splits into two clusters, and so on until each "cluster" contains only a single item. Based on the problem under consideration, this process can be stopped at any point to create meaningful clusters.

Common uses in healthcare:

- **Risk adjustment.** Risk adjustment is a crucial analytical tool for many applications in healthcare, from clinical-outcomes studies to determination of reimbursement for patients in capitated care delivery programs (such as Medicare Advantage). The problem is that, when calculating the impacts of different activities on outcomes, the baseline level of illness for each patient needs to be computed to create a consistent baseline for comparison of methods across all patients. Clustering can be used to group patients into meaningful groups of similar comorbidities or risks.
- **Patients like mine.** In the care of complex or rare disease, it is valuable to understand how different treatments have worked on other patients in similar situations, but this information is only useful if the past patients are similar enough to the current patient to have predictive value. Clustering can be very useful to identify the most similar patients.
- **Population health management and chronic disease management.** Current population health management methods rely on identifying broad groups of patients for whom interventions can be implemented to improve outcomes in general. Clustering is very powerful for finding these groups of patients.

7.9.10 Text Understanding

What does this text mean?
Example: Is this a properly documented diagnosis of diabetes with peripheral neuropathy?

- **Natural Language Processing (NLP).** Natural language processing includes an extensive toolkit of different text processing tools and steps, combined with the goal of understanding the meaning or practical implications of a piece of text. In clinical text analysis, text understanding depends on being able to recognize distinctions among diagnoses attributed to a patient, those mentioned in a differential diagnosis, family history, and negations ("patient does not have diabetes mellitus"). There are a number of open-source tools for general NLP and for clinical NLP that can be incorporated into application development. Full NLP analysis of text can be very computationally expensive, both in terms of CPU cycles and memory required.

- **Text mining.** Text mining is the application of extensive dictionaries of terms to identify occurrences of key terms in text such as clinical notes, consult letters, and discharge summaries. Text mining has the advantage of being able to recognize vast variations in terminology, including abbreviations, misspellings, regional variations in usage, and transcription errors from scanned documents (often using optical character recognition, or OCR). Text-mining methods are often augmented with specific NLP tools to help understand the context of the terms that are identified in the text. For example, negation detection can be combined with search for diagnoses to help understand the difference between a positive statement of a diagnosis and a negative statement that a diagnosis does not apply.
- **Deep learning.** As described above, deep learning has the ability to learn key features of data without explicit programming. In the case of text understanding, deep learning has begun to show value for identifying complex ideas in text and interpreting the implications of their context.

Common uses in healthcare: Studies show that structured data in healthcare can be deeply flawed. For example, structured problem lists consistently suffer from extensive false negatives and false positives, even for impactful conditions such as heart failure. As a result, clinical text is one of the most reliable sources of usable information in healthcare, and the number of healthcare systems using text analytics as part of their reporting, decision support, and care optimization efforts is growing rapidly. Examples of applications include:

- Identify conditions that have been addressed in face-to-face encounters but not submitted to Medicare Advantage for capitated payment.
- Automated coding of typed or dictated encounter notes.
- Extraction of key findings in radiology reports to correlate with information in the EHR.
- Identification of medications and other key findings for inclusion in risk-prediction models.
- Automated chart review to identify care, such as annual diabetic foot exam, which is routinely performed without being separately coded. These reviews can directly impact reported performance measures.
- Mapping of patient care history, across multiple provider organizations, without requiring access to data sets from all providers.

Figure 7.4 shows how machine can be combined with other types of analytics to solve a large swath of business problems.

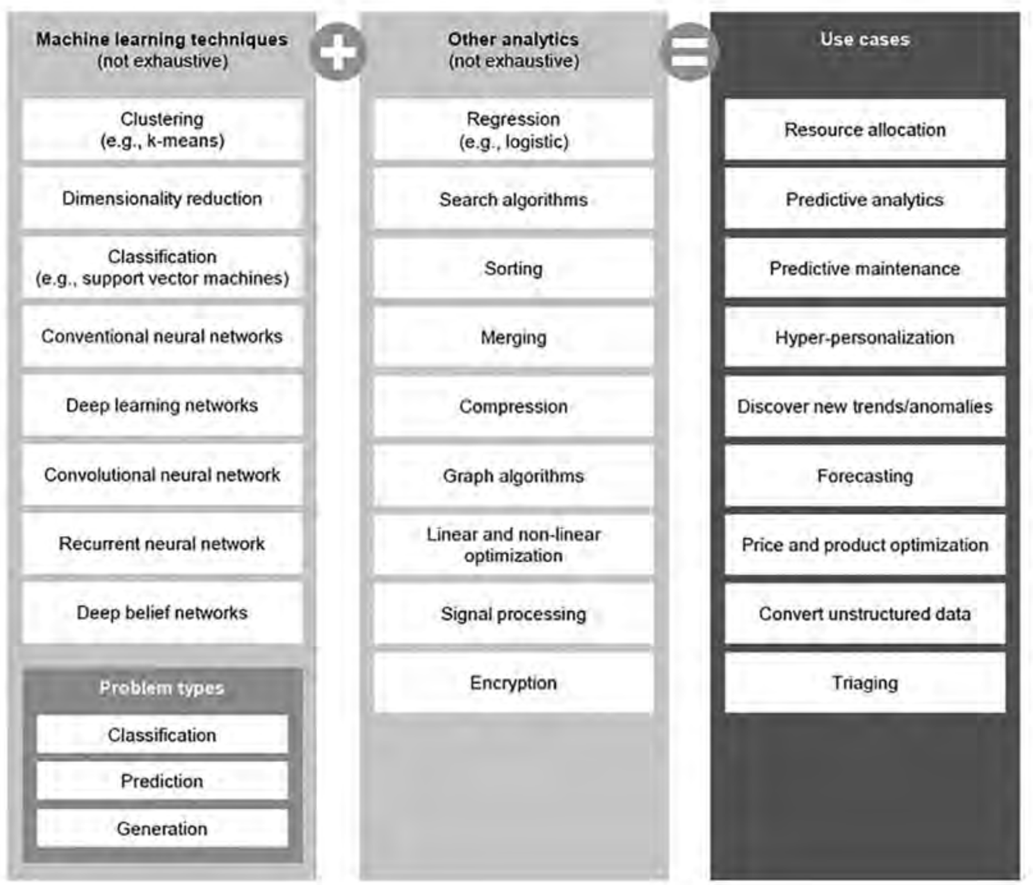


Figure 7.4 Machine learning can be combined with other types of analytics to solve a large swath of business problems. (*Source:* Exhibit from "The Age of Analytics: Competing in a Data-Driven World," December 2016, McKinsey Global Institute, www.mckinsey.com. Copyright © 2016 McKinsey&Company. All rights reserved. Reprinted by permission.)

7.10 Machine Learning and the Contextually Intelligent Agent

Machine learning will begin to realize its full potential in healthcare when contextually intelligent agents (CIAs) are put into widespread use. A CIA is a system that can interact directly with a human, via spoken or written communication, and that can understand context to identify what's important in a given situation. CIAs are a first, and crucial, stop on the journey to artificial intelligence. (See Figure 7.5).

Why is context so important? As we just described, most machine-learning algorithms have been created to answer a specific question. How can I compare two patients? Is this a heart failure patient? When will this patient convert from pre-diabetes to full-blown diabetes? But for any given situation, there are many, many such questions that could be reasonably asked, and then answered with a machine-learning algorithm. Intelligence depends, in part, on the ability to know which question to ask at any given time, based on the context of the situation.

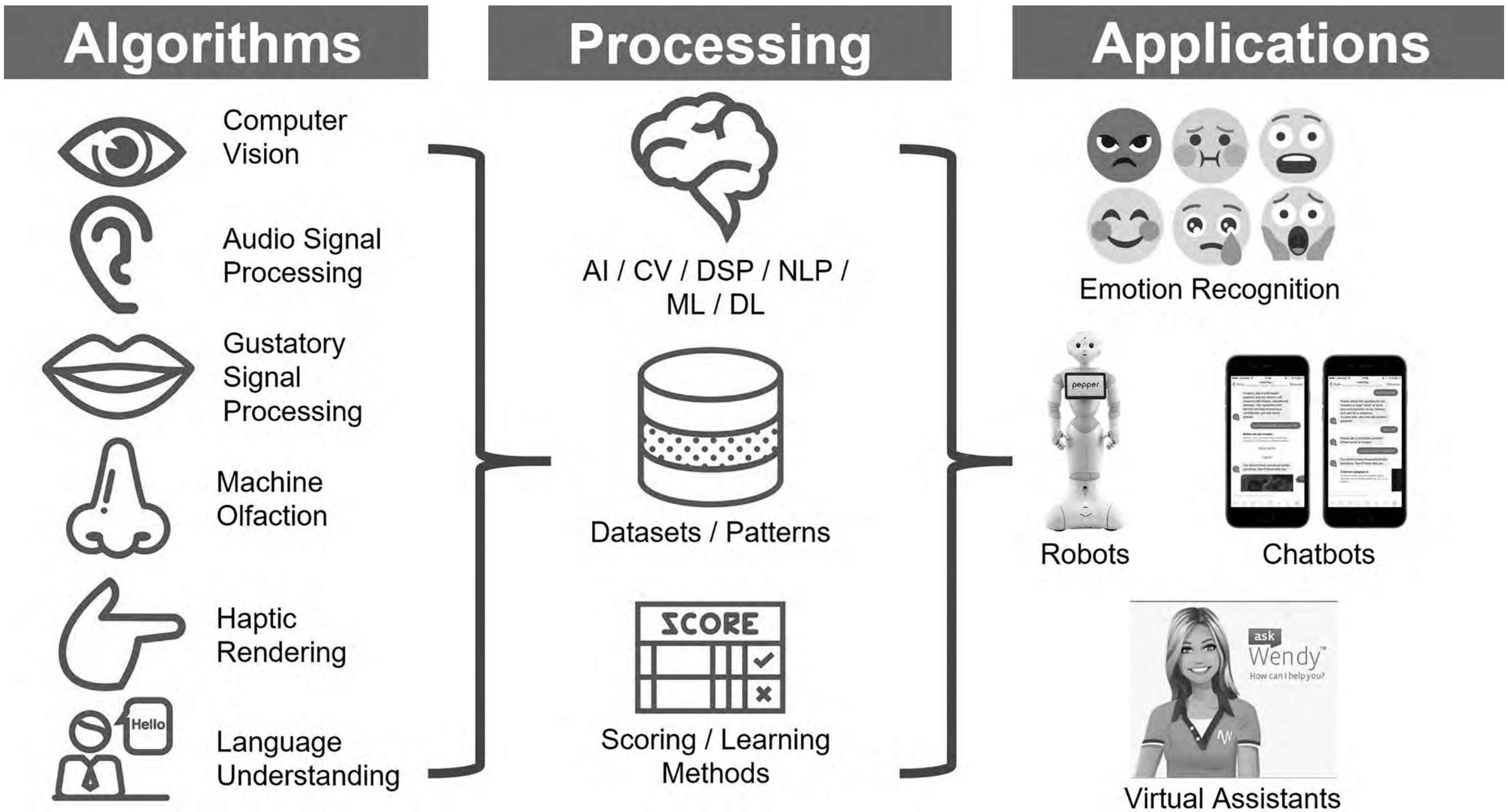


Figure 7.5 Learning algorithms, processing, and CIAs in healthcare. (Used with permission of Catalaize, Nardo Manaloto, and Wen Dombrowski MD.)

Who is asking the question? What has recently changed? Where is the question being asked? We refer to systems that are sufficiently contextually aware to be able to ask (and then answer) the right questions as CIAs.

CIAs are beginning to have an impact in healthcare. Chatbots are effectively communicating with patients to help them find an effective disease management intervention or to coach them through physical therapy following a surgical procedure. Or consider the portable ultrasound, a revolutionary technology that is being actively developed right now. To have non-invasive, highly actionable ultrasound information available in the field, where it can be most valuable, is a huge advance, but the challenge is that it has traditionally required an enormous amount of training for an ultrasound technician to be able to collect clinically meaningful images. How can a relatively untrained user in the field collect clinically useful ultrasound images? The answer is that the portable ultrasound comes with a CIA that can guide the user to collect good data given the circumstances that are driving the need for the imaging. Such contextually aware applications will become increasingly important in healthcare and will drive the value of machine-learning applications.

We would argue that, just as machine learning was a necessary technology to extract value from big data, the widespread availability of CIAs will be required for machine learning to realize its full value in healthcare. Not only do algorithms need to give the right answers, but they need to ask the right questions: questions that are sensitive to the current clinical situation, and which fit neatly into a clinical workflow rather than disrupting it. CIAs, which are appearing in many areas of healthcare, from payor disease management programs to patient wellness and diagnostic imaging in the field, are the key enabling technology to allow this transformation to happen.

7.11 Some Best Practices for Successful Machine Learning

7.11.1 Ask a Specific Question

Your machine-learning algorithm should answer a very specific question that tells you something you need to know and that can be answered appropriately by the data you have access to. The best first question is something you already know the answer to, so that you have a reference and some intuition to compare your results with. Remember: you are solving a business problem, not a math problem. Ask yourself, "What valuable action will be taken as the result of my analytics?"

Analytics and artificial intelligence systems come in two flavors: (1) knowledge management systems that interpret questions and provide information to answer these questions, and (2) very targeted quantitative systems designed to provide information for a specific use case. Don't try to build both types of system in a single effort.

7.11.2 Start Simple

This is true for model selection and the data you consider using for your analysis. You want your results to be robust, so less model complexity and fewer parameters are always beneficial. Regarding data, don't start by building a huge data lake with every kind of data you could possibly get your hands on. Instead, start with the minimal set of data that could get you to a good result.

7.11.3 Try Many Algorithms

Most machine-learning toolkits support multiple algorithms. Try a few to see how they work. This allows you to find the best tool for the job. Also, if one classifier works incredibly well and another doesn't seem to work well at all, be cautious. You may have an overfitting situation, which means you won't really have much predictive power. You may also want to combine methods: use deep learning to extract features from unstructured data and then use these features, along with others, in a classical machine-learning algorithm to get interesting results.

Remember that data is more important than the exact algorithm you use. More training data is always desirable. In addition, for classical machine-learning applications, the better your features, the better your performance will be.

7.11.4 Treat Your Data with Suspicion

Look at your data, dig into its details, look for correlations, suspicious gaps, systematic biases, errors, and flaws. Use statistics and visualizations here. Text has transcription errors, misspellings, and abbreviations. These challenges often exist for structured data as well: you will find that data is recorded inconsistently both across your data set and even within a single field.

7.11.5 Normalize Your Inputs

Machine-learning algorithms can perform poorly if there are large differences in scale between different features.

7.11.6 Validate Your Model

Separate your data into training, test, and validation sets, or if you are using K-fold cross validation, at least hold out a validation set. You need to keep some powder dry for most applications. Also, be aware of biases in your split. Remember: there is no such thing as a random set of data, only a random process to generate data. If you randomly flip six coins and they all come up heads, that's not going to be a very good validation set.

7.11.7 Focus on Data Fidelity But Ensure Quality of Training Data

Data fidelity is more appropriate for machine-learning systems than data quality for the reasons discussed in Section 2.4.2. For supervised learning algorithms, you will want to look closely at your training data. Does it cover all the use cases? Is it biased in some way? For example, did multiple humans create it? Can you see biases or differences among different folks?

This is particularly challenging in healthcare, where unstructured data is critical and source data comes from multiple silos. Extra effort in developing a high-quality training set will pay major dividends and will improve fidelity. Because of the variation in how information is represented in different healthcare settings, the more diverse the sources of data you use in your training set, the more transferable your results will be.

7.11.8 Set Up a Feedback Loop

Think through how you will use the output errors of your machine-learning system to improve it. Downstream users can provide feedback on when your algorithm got it wrong. How are you capturing this feedback so you can bring it back into training?

Note: This is great for false positives, but can miss false negatives, so you will want to pay special attention to false negatives as you train and use this experience to help you find missed results in production data review so you can include them in your next round of training.

7.11.9 Healthcare Doesn't Trust Black Boxes

Some machine-learning methods are more transparent than others. Clustering, topic modeling, and recommender systems tend to be easy for humans to interpret, because they create groupings of concepts that humans can associate with known influences. Linear regression can tell you how important each feature is to the final output. This is true to a lesser extent with linear classifiers. Random forests are difficult to interpret. Deep learning is truly a black box, with very little transparency to what is important in the decision-making process.

Note that there is a lot of research in this area of machine learning, so better tools for helping us understand the decision process are coming. In fact, some third-party healthcare analytics providers have instrumented interesting explanatory tools for their machine-learning algorithms.

7.11.10 Correlation Is Not Causation

It's easy convince yourself that two factors that move together imply that one causes the other. Just remember that in many cases there is a hidden factor that could be causing both factors to move together.

7.11.11 Monitor Ongoing Performance

How will you monitor the performance of your algorithm on an ongoing basis? Data drifts and systems evolve. You can do this manually by spot checking your results against the incoming data, and you can monitor data and algorithm statistics with a dashboard. Simple moving averages can tell you a lot.

7.11.12 Keep Track Of Your Model Changes

Always track the revision of your model and report it with your results. As you improve different parts of your data analytics pipeline, you will want to go back and re-analyze data. Recording which model was used at which time helps you understand what to recalculate.

7.11.13 Don't be Fooled by "Accuracy"

If you're looking for a rare event that only happens 1% of the time, and you never actually find it, you can report your accuracy as 99%. Obviously, that's meaningless. Instead, figure out before you start your project what precision and recall your application requires to be useful. Build your application to these metrics.

7.12 Conclusion

In conclusion, let us look at some next steps for machine learning and AI in healthcare—and what it means to professional practice, personal skill sets/knowledge, and jobs.

1. The current state of the industry, as evidenced by rapid recent progress and increasing investments, is promising. However, we haven't reached the Promised Land yet. We will get closer to it when the predictions and data-driven insights from machine learning are connected to contextually intelligent agents (CIAs), as described in the preceding section. CIAs and applications (see Figure 7.5) are necessary to complete the "data to action" or "data to behavior modification" workflow loops. They will be critical to the "mainstreaming" of machine learning within each HC organization and by individuals—patients, providers, and consumers included.
2. As you can expect, machine learning and AI in healthcare must deal with or account for the foundational five V's of big data in healthcare, as described in Chapter 2. Successful advanced analytics (predictive or prescriptive analytics) and CIAs will "need to be able to easily integrate more data sources, harness machine learning and advanced technology for faster, more sophisticated analyses, and extract insights that will improve business performance."[49]
3. Labeled data in healthcare remains a challenge and a time-consuming effort. While making more data available to machine-learning systems is always helpful, in healthcare setting that data must be prepared and labeled in the context of the source and intended use. Context-based labeling is essential to ensure the relevance and validity of learning, prediction, and action.
4. Feedback loops are essential to any healthcare machine-learning system. Period. Feedback loops can be of two types:
   - Explicit annotation by human experts
   - Implicit inferencing from downstream use of the data as evidenced by the behavior of both
     - o contextually intelligent agents, and
     - o individuals who are the recipients of the recommendations/actions suggested by the agent/application

[49] "How Analytics and Machine Learning Help Organizations Reap Competitive Advantage." (2016, December.) *MIT Technology Review.* Available at https://s3.amazonaws.com/files.technologyreview.com/whitepapers/Google-Analytics-Machine-Learning.pdf. p. 3.

5. The topic of data fidelity (as opposed to just data quality, as discussed in Sections 2.2.2 and 2.4.2 in Chapter 2) remains an important one in deep learning today. Some practitioners believe that deep-learning systems require the highest data quality as inputs for the predictions/results to be relevant and useful in healthcare. Other practitioners believe that imposing more stringent data quality requires more up-front investments and longer time to market. According to this school of thought, the advantages of deep-learning black boxes are proportionately reduced with increased attempts at enforcing transparency and management of uncertainty in hidden layers of a deep-learning system.

   The reason for this debate is that current deep-learning experiences don't provide needed transparency and do a sub-optimal job of determining and managing uncertainty in hidden layers or can't always identify blind spots between input and output layers. We agree that more transparent management of uncertainty and weights will allow us to do more deep learning on healthcare data of lower-data quality than is currently required.

   Given the current state of deep-learning architectures, we also think this debate is yet to be fully settled in healthcare. As a result, we don't yet have validated and settled best practices for this topic. However, given the coming importance of contextual usage-driven intelligent/applications for deep learning, we strongly recommend you use data fidelity (as defined in Section 2.2.2 of Chapter 2) instead of data quality in the design and deployment of deep-learning systems, both in the interim and for the future.
6. Questions for readers:
   - How are you approaching data fidelity in deep learning?
   - Do you think we're close to settling this debate or will that happen only when layers are "unhidden"?
   - What are the estimated added costs (or conversely, savings) of going from black-box to white-box deep learning or improving data fidelity?
7. In an Analytics 3.0 environment, the roles of data integration and MDM remain as critical as they do in the Analytics 1.0 and 2.0 environments. Analytics and workflows based on machine learning and CIAs must be integrated with all relevant data—both little and big—across silos in order to benefit the healthcare enterprise and users. Your EDW will continue to play a key role in supporting newer processes and technologies in an Analytics 3.0 world.
8. Questions on ethics and privacy as they relate to machine learning (and AI) are as relevant today (if not more so) as in the past. Current AI systems are not designed to account for the morality of learning algorithms and machine learning. It is useful to point out an important distinction between human and machine morality: "The moral constraints to which we are subject in our dealings with contemporary AI systems are all grounded in our responsibilities to other beings, such as our fellow humans, not in any duties to the systems themselves."[50]

[50] Bostrom, N. and Yudkowsky, E. "The Ethics of Artificial Intelligence." (n.d.). Machine Intelligence Research Institute. Available at https://intelligence.org/files/EthicsofAI.pdf. p. 7.

While it's beyond the scope of this chapter to provide answers to the various questions already being raised, we would like to suggest the following topics for more discussion and research.

a. How do we forward the knowledge coming out of machine learning and AI to the patient in a transparent and ethical way? How do we establish provenannce of insights being shared with the patient and provider?
b. Ethics of the human participant as related to providing inputs or acting on outputs.
c. Robo-ethics: ethics built into the machine-learning system by design or during learning.
d. Transparency around access to and inspection of the machine-learning system.
e. Reliability of predictions and validated performance.
f. Clear demarcation or sharing of human and machine-learning/CIA responsibilities when failure happens.
g. Legal, privacy, and innovation (patents, copyrights, etc.) considerations.
h. As you can guess, points a–g above will also create new discussions on the ethics of AI as related to public policy, reimbursements, population health management, SDoH, and access/cyber security at the global and individual levels.

9. Jobs: While applied machine learning and AI are here to stay/thrive, the impact on healthcare jobs will mostly be positive (helping humans do their tasks or creating more jobs) in the immediate- to mid-term. As a result, we will see increased augmentation of human tasks in healthcare—and not the overblown wholescale replacement of physicians, surgeons, radiologists, nurses, CXOs, or IT geeks as often portrayed in the popular press.

   However, we do not recommend complacency. In the mid- to long-term, we fully expect to see more administrative roles and even some specialties to be eclipsed or replaced by machine learning and AI (and in certain geographies and organizations at an accelerated pace). So, how do readers prepare for coming changes in healthcare jobs?

   a. Understand what's available and coming; we hope this book has helped you get started.
   b. Do projects to investigate and be prepared for emerging technologies in this space.
   c. Leverage your existing skills to drive question determination, feature definition, labeling, data integration, feedback loops, and promotion of data exchange and use.
   d. Leverage intelligent bots, agents, etc. and try things out via smartphones, personal wellness applications, CIAs.