AI: Executive Briefing
by Simon Allardice

Start Course

Bookmark Add to Channel Download Course

AI: Executive Briefing

Introduction

This is a short course about artificial intelligence, or AI, and we begin with the problem. The problem with AI is that everybody knows something about this already. The idea of artificial intelligence is part of our culture. It's in our movies, our TV, our books, our games. Even in the earliest days of feature films, we had Metropolis nearly 100 years ago; 2001: A Space Odyssey over 50 years ago; or Blade Runner; WarGames; The Terminator; The Matrix; I, Robot; Ex Machina. You've been hearing about intelligent computers for decades, probably all your life, and--spoiler alert--where AI always ends up wanting to kill everybody. Now it's easy to think, yeah, yeah, yeah, but that's fiction. I realize that. But you know, when we hear and see similar cliches again and again and again, we should recognize we have probably all unconsciously picked up some beliefs and assumptions about artificial intelligence. And I don't mean that we've all agreed on the inevitable dystopian nightmare future, but more generally, just that we've absorbed some ideas and assumptions about what artificial intelligence means. What do we think AI is capable of? What we think it's not capable of? And most people also have a few assumptions about how complicated this is, or how much of a computer nerd they have to be to even understand it. But here's the thing. If you've been, let's just say, just aware of artificial intelligence for a while, you might also have a sense you've been hearing, a lot more about it recently, and not just in fiction,

but seeing artificial intelligence mentioned in more business publications or seeing more conferences about it, more websites that mention it, more products that tout AI as a feature. And yes, it has become much more important recently. We've seen an explosion of applied, pragmatic, real-world applications. Artificial intelligence is being used in healthcare, being used in transportation, being used in finance, in marketing, in education, in customer support. It is everywhere. And over the next few minutes I'm going to take you on a tour of the most important ideas about artificial intelligence, different ways it can be implemented, and we'll go over a lot of the vocabulary. We'll explore ideas like machine learning, deep learning, neural networks, natural language processing, we'll talk about deepfakes and chatbots and more on. And we'll focus on real world applications, how AI is actually being used and implemented right now. I'll cover a lot of the basic fundamental questions people often have when they get started here, even things as simple as, so what's the difference between artificial intelligence and just a very complex computer program? Is there a difference? And if AI has been around for decades, and it has, why has it recently become such a big deal? What happened? But we'll begin with probably the most common question. What exactly is artificial intelligence? What does that even mean? Is there a simple, straightforward definition we could learn and internalize and from that point on, always be able to say, okay, that is artificial intelligence and that is not. Another spoiler alert, the answer's no. There is no single agreed-upon definition that will always let us make that distinction. However, the question, what is artificial intelligence, is still an excellent question and a great place to begin. I'm Simon Allardice. Welcome to the executive briefing on AI.

Defining Artificial Intelligence

One of the reasons I enjoy talking about AI is it's simultaneously the most innovative, current, up to the minute, and groundbreaking technology you can imagine, which also has a very long and substantial history. This is older than anything else I teach. I have a few textbooks here on artificial intelligence. This 1 is 30 years old, this 1 is 40 years old, this 1 is 55 years old. I can't exactly go out and find a 55-year-old book about web development, or block chain, or IOS, but artificial intelligence has been a serious field of study since the mid-1950s, and from those very earliest days, they never shied away from admitting this phrase, artificial intelligence, is not simple to define and nail down, and not because it's difficult, we understand what these two words mean, but because, and here I'm going to use a very unusual for me sport analogy, with AI the goalposts keep moving. This is a paper on artificial intelligence from 1958 written by Marvin Minsky, 1 of the major initial researchers in the field, he cofounded the AI lab at MIT. I'll read two sentences. It would not be useful to lay down any absolute definition of intelligence or intelligent behavior for

our goals in trying to design thinking machines are constantly changing in relation to our ever-increasing resources in this area. As a little perspective, when he wrote these words, there would still be another 40 years to go before Google would even exist, but through the entire history of AI, the entire history of computing, we've always had these goals, these aspirations that we've described as requiring some kind of intelligence like can a computer ever play chess at a grand master level, or will a computer ever recognize faces, or could a computer ever understand spoken commands? And as soon as we've ever figured out how to do that thing, the general response becomes okay, but that's not really intelligence, is it? The computer isn't actually thinking, it's just executing a formula, it's just a computation. There is even a name for this. We call this the AI effect that with many technologies we now use every day they begin as some kind of artificial intelligence research, but as soon as any part of it became implemented, practical, and widespread, nobody will call it AI anymore. There are tasks you could do on your phone that a few years ago would have been considered remarkable, incredible implementations of artificial intelligence, As just one small example, we have facial recognition that so fast I could turn myself into a 3D cartoon in real time. If I presented this at an artificial intelligence conference 10 years ago, I would have been the star of the show. Now we don't even use the term AI to describe it, it's just something we can do on a phone if we even care. So we're not going to worry about having the perfect unchangeable definition of AI, but sure, we need some kind of definition, even if just too broadly categorize what are the kinds of things we typically consider artificial intelligence versus regular conventional computer programming? And I'm going to say that instead of going from the top down, you know where first we must precisely define what is artificial intelligence and only then can we start talking about examples, it's actually better to build our understanding of AI from the bottom up. Let's explore some examples. Let's explore the approaches people have used to get a computer to have some kind of behavior we might call intelligent and that's how we'll get to a worthwhile understanding of this. Though it is worth pointing out, there are people who work in this field who really don't like the phrase artificial intelligence at all.

Booms & Busts of AI

If you've been in the business world for a long time, let's say several decades, or you work with other people who have, there is a perception about AI you may need to deal with, which is didn't we already try this back in the 1980s and figure out it was a total waste of time? And yes, there have been previous waves of AI, periods of massive enthusiasm about it, followed by years of deep skepticism. I started programming in the mid 1980s, and there were a lot of AI vendors springing up at that time, often selling what were called expert systems. There were combinations

of software and even dedicated hardware, there were systems like XCON and Intellect, we had languages like OPS5. There were companies that specialized in this. We had Symbolics and IntelliCorp, and even an artificial intelligence corporation. Now, if you're thinking I've never heard of any of those, yeah, I know, because to cut a very long story short, while that period was a wonderful time for most of the computing industry who just kept on growing from strength to strength, the entire AI segment of it just crashed and burned. Most of the companies went under or pivoted away, and AI was generally regarded as a massive failure, particularly in the business world. It had cost a ton of money, it had promised the world, and it hadn't delivered. And the impact was big enough that even the term artificial intelligence had a stigma about it. And for many years, if you wanted investment or funding, the last thing you wanted to say is you had an artificial intelligence project. Nobody wanted to know. This period is called an AI winter, and there's been more than one AI winter, a long, dark time of no money for it and no interest in it. And there are companies to this day that avoid using the artificial intelligence term because of all the baggage it has.

AI Cliches: Good and Bad

More than anything else I teach, artificial intelligence is full of cliches, things that everybody kind of knows about this, and we have to identify and get past some of them. And I don't just mean the old AI will inevitably try and kill us all story because some people think that risk is something to be taken very seriously. But there're even visual cliches about this. For example, have you noticed that in our current culture, artificial intelligence has a color? Apparently, we all think AI is blue. Now if you're thinking, Simon, what are you talking about? What do you mean AI has a color? Well, let me prove it. I'm going to do an image search on AI. We'll gather different pictures from all over the web associated with this word, thousands of images created by countless different artists and designers, but it's like most of them thought they were only allowed to represent AI in one of two ways, a wireframe image of a brain, which must be made of nodes, or a robot, but not just any robot. It must have a white plastic face and, for bonus points, exposed head wiring. And notice, almost everything is blue. And this isn't just true for images on the web. It's true for book covers, and you'll see the same thing in presentations. It's like all the graphic designers had a meeting and agreed, Well, AI is blue, obviously. I'm poking a bit of fun at this, but there is a point. I want you to start to notice these cliches. So the next time you see a book or an article about AI and the first thing you get is blue image of robot with creepy doll face, you think, Okay, seen this before. What other AI cliches am I about to get? And the thing is, some of the cliches about AI are actually useful. They give us starting points. They give us ideas to talk about.

But some of them aren't, and I actually think these images are slightly harmful. They're slightly misleading because when we're talking about real-world, applied business-focused artificial intelligence, there're two things we're usually not trying to do, model a brain and create a robot. Now let me be clear. Both of these ideas, brain modeling and robotics, are active areas of AI research, and there was a time when a lot of AI research took this neurobiological focus. And we thought that success would only come from deeply understanding the physical structure of a biological brain and creating a software version of all the neurons and synapses. But while that is one approach to AI, it is not the only approach. And many of the practical successes of recent years have taken completely different strategies. They're often more to do with statistics than biology. And if you read articles and books and presentations about current applied business-focused AI, to be honest, it's mostly about the practical techniques that have nothing to do with brain modeling or robotics. But they bombard you with supporting visuals that imply this is all about brain modeling and robotics. But I said there were some useful cliches. And, next, I want to talk about perhaps the most common one you'll find in movies and TV and novels, the idea of an AI that operates at the level of a person.

General vs. Narrow AI

And most fiction about artificial intelligence, it's used to generate an entire personality with a wide range of abilities. Sometimes that fictional AI has a body, whether humanoid or a machine. Sometimes the AI is running on a computer behind the scenes, and the interaction is with microphones and speakers and screens, but in either case where that fictional AI is typically capable of complex, multilayered conversations and behaviors similar to possibly higher than the level of a human, and where it's not just capable of a few specific and narrow tasks that is generally intelligent, it's capable of many skills and capable of learning new skills by itself. And this is what's known as artificial general intelligence, orAG,I also called strong AI or full AI. Now, just to be clear, this is still hypothetical, it's theoretical. We don't have AGI yet. It's not even the goal or the expected outcome of a lot of what we do, but we do have the other kind of AI. Some call it a weak AI. I prefer the term narrow AI because weak is a word that implies failure, implies deficiency, and we're talking about AI that is often extremely powerful, something that can typically outperform a person, but only in one specific narrow ability. So we already have AI that can play chess better than anyone in the world. We have AI that's great at spam detection, getting better. We have AI that's wonderful at recognizing faces. We have AI that can scour massive amounts of data like search engines or recommendation systems and provide insights about that data, but with narrow AI, we understand that just because a computer program might

be superb at detecting a spam email or recognizing an abnormal heart rhythm, it doesn't mean that same program can also play chess, compose a sonnet, and recommend the next TV show to watch. And that's okay. For the most part, we're not actually looking for general intelligence from it. What we want is specialized intelligence, narrow intelligence because we're trying to use AI to accomplish or understand something in particular. We want a program that is task specific. Now, if you're wondering, so is anyone actually trying to develop artificial general intelligence? Oh yes. There is billions being invested in this. There are lots of organizations working on this right now, but I'll name just two. For example, DeepMind, and I'll quickly point out DeepMind should not be confused with Deep Blue, which was a chess playing AI developed by IBM. Now DeepMind is a company now owned by Google and is very specific that their aim is to build advanced AI, sometimes known as artificial general intelligence or AGI. Another example is OpenAI, that's a research lab that has billions in investment from companies like Microsoft and Tesla, and their stated mission is discovering and enacting the path to safe artificial general intelligence. So, yes, there is a lot of research and development on this, but with an executive briefing course, we're less focused on R and D and more about practical applied AI. So what can we actually do with this right now? Now let's step into one of the ways we can implement a narrow AI today using machine learning.

Introducing Machine Learning

Over the last few years, the most widespread, successful, and practical approach for most companies implementing AI has been with machine learning. And the success of it has even lead to some confusion about the terms. Machine learning and artificial intelligence are often mentioned together so much that they can seem like they mean the same thing. And they don't. As you're probably beginning to realize, artificial intelligence is itself a rather broad and sometimes vague term that encompasses, well, all the various approaches we could use to make a computer program do something smart, and machine learning is a set of specific techniques for achieving that. It's not the only way to implement AI. Machine learning, or ML, is a type of AI. It's a subset of it. So all machine learning is artificial intelligence. But artificial intelligence includes other things than just machine learning. Now, in a regular, conventional computer program situation, we would try and get the computer to do what we want by explicitly programming a bunch of rules and very specific instructions into it, lots of commands and conditions and if/then statements, if this number is higher than that number, then add these two things together, if this date is less than that other date, then generate error message, if the position of the asteroid on the screen is the same as the position of the spaceship on the screen, then play boom sound effect and display

Game Over. Now, make no mistake, conventional rule-based computer programming is fantastic when you know what your rules are and you can define them. But when we use machine learning, we try and get to the results that we want, not by providing a bunch of rules to follow, but by providing examples to learn from. Let's say I was trying to create a computer program that could take a small digitized image and recognize handwritten numbers, handwritten digits in that image. Well, let's make it even simpler. I just want to write a program that could recognize the number 3. If I take the conventional approach, I might try writing a program that would look at each individual pixel in that image, one by one, figure out if it was light or dark, and then try and think of all the different ways people might ever write a number 3 in this grid. Sometimes it would be straight. Sometimes it would lean to the left or the right. Sometimes the top loop is smaller than the bottom one. Figure out all the rules, go step by step in trying to envision every single eventuality. And to be clear, this would be hard, tedious work and, with this particular problem, very easy to get it wrong. With machine learning, we begin with actual data. In this case, that might mean perhaps 1000 or 2000 different images of a handwritten number 3, and we would use the data itself to build our program. We would take our data, our examples, and we would feed them into a machine learning algorithm, also called a learner, in a process which is called training a model. And we often provide both positive and negative examples. So here's a bunch of examples of what we are looking for. Here is a bunch of examples of what we're not looking for. And after we've trained a model, we could then expect to be able to take a new piece of input, a new handwritten image, and then have that model tell us, Does this new image looked like a 3 or not? It doesn't mean the computer truly understands what 3 represents and where it fits or even how to write one. It doesn't. We've just provided enough examples, it can recognize the characteristics of that thing. And this is one of the tasks machine learning is very, very good at, classification. But let me be clear. Classification isn't just about images. It works on any kind of data. We could take the same concept into things like, Is an incoming email spam or not spam? Or if we have a bunch of web site activity data, could we use it to figure out the typical characteristics of high value customer versus low value customer and then be able to recognize earlier on when that's happening? But there is a catch. Machine learning can only work if you have data to learn from, and not just a little bit. You want a lot of data and you want good quality data. It's why there's been such a focus on data collection and big data in the business world over the last few years. If you don't have good data, you have nothing for your machine learning algorithms to learn from. Now a common question at this point is, If we do this by taking some data and feeding it into a machine learning algorithm, what is that? Where does that machine learning algorithm come from? Well, let me show you.

Machine Learning Platforms & Frameworks

Machine learning requires data, and it requires machine learning algorithms that can learn from that data. And when you're new to this, you might wonder, okay, so did the software developers in my organization have to write these machine learning algorithms? No, the majority, the overwhelmingly vast majority of companies do not need to write and never need to write machine learning algorithms. Instead, we can find a machine learning algorithm somebody else has already written, and we can feed our data into it. But Simon, you say, you're telling me we just pick some machine learning algorithm off the shelf and we can just shove our data into it? That's a bit oversimplified, but yes, the same way that when we want a database, we would typically choose an existing database management system. Or if we want to make a website, choose an existing web framework. That applies here, too. When you're using machine learning, you're typically going to choose an existing machine learning platform or machine learning framework. And there's a lot of them, both commercial and open source. It's extremely common to use a cloud-based, hosted machine learning platform. For example, Microsoft Azure has one of these, Google Cloud Platform has one, as does Amazon AWS, as does Alibaba, IBM Cloud. They all provide the frameworks and a set of tools to help you do all this. And they have multiple machine learning algorithms we can actually pick and choose from and apply to our data, depending on what kind of task we're trying to do. We don't need to write these machine learning algorithms ourselves. And this has been the biggest enabler for the huge growth of applied AI in the last few years. It's this combination of having these pre-prepared machine learning frameworks ready to go, together with the computing power you need to run them in the cloud, where you only have to pay for what you use. And it's reduced the barrier to entry. It's made it so very simple to start doing this. But even though we may not write these machine learning algorithms ourselves, there's still a lot of work to do. You can't just feed all of your corporate data into some machine learning algorithm and expect it to understand it all for you. And the most important part of this entire machine learning process is preparing the data, understanding what it is we're looking for, and then filtering that data, cleaning it, labeling it. If your data is full of garbage, full of invalid values and missing values and conflicting information, it doesn't matter how good the machine learning algorithms are. As ever in programming, you put garbage in, you get garbage out. If you have a team of people working on machine learning projects, this is one of the responsibilities of the data scientist role, transforming the data into something meaningful, something usable. And after that, choosing and applying the different machine learning algorithms. And this is another part that takes time and attention and, most importantly, understanding what it is you're trying to find, what you're trying to do. Most data scientists will tell you it's not a single task; it's a process.

They'll often experiment with several different machine learning algorithms and test them to see which ones are giving the best results for the current problem. In an introduction like this, we can only scratch the surface. If you're interested in learning more about machine learning, I have an executive briefing course just on that here at Pluralsight.

Deep Learning and Neural Networks

Earlier, I showed that visual cliche of the wire-frame brain, and I said a lot of the recent successes with AI used other approaches than modeling a brain. And when you're using any of the popular machine learning frameworks, you can choose between a lot of different prewritten algorithms. They have names like Support Vector Machines, and Decision Trees, and Logistic Regression. I'm not expecting that you know the difference between all of these. I'm just naming a few examples, but many of them are based on ideas, and statistics, and probability. They have nothing to do with the brain. However, just as machine learning is a subset of AI, there is also a subset of machine learning called deep learning, which uses algorithms that are more to do with the brain. These algorithms are called neural networks or artificial neural networks to distinguish them from the real biological neural networks like the ones we have inside our skulls. Now it'd be an exaggeration to say these algorithms are actually modeling the brain. It's more accurate to say they're inspired by a teeny little part of it, the idea that our brains have all these neurons and synaptic connections, neurons with multiple connections to other neurons. And in an artificial neural network, that idea is somewhat modeled. There are simulated neurons, individual nodes that could be connected to and send messages to multiple other simulated neurons. An artificial neural network is made of multiple layers. There's always an input layer, and an output layer, and typically at least one hidden layer in between them. When it's more than one hidden layer, we call it a deep neural network, which is where the deep and deep learning comes from. And each node, each neuron is capable of performing some small function, some small computation as data flows through this. Now this is still machine learning, which means we have to provide examples of the results we're trying to get to. We provide examples of what success looks like. So a neural network, when we're training it, can take some input, and it will start passing it through the different layers, all the different nodes and doing this again, and again, and again. And as operations are performed on each piece of data, the output from the different nodes will lead to results, either closer to what we're trying to get to, in which case, those outputs become weighted, and the connections become strengthened, or the data will skew further away from what we want, and the importance of those connections becomes minimized. If this sounds a little abstract and high level, it is because there really isn't much further I can go into neural

networks without starting to say things like stochastic gradient descent and back propagation, and this isn't that course. So instead, let's focus on this idea. What's it good at? Well, deep learning, using these deep neural networks, that's what's giving us the best results in complex situations like image processing, facial recognition, speech recognition, and complex game playing. It's capable of very impressive, very profound results. But there are a couple of downsides. Training a deep learning model can be much more computationally intensive and take much longer than taking some of the other simpler machine learning methods. And part of that reason is that deep learning also typically requires a lot more data to train properly. Now it's always difficult to say exactly how much data do you need to train a model. It's very situation dependent, but certainly it wouldn't be unusual to expect to have tens of thousands or even hundreds of thousands of examples in order to train a deep learning model well, but once that model is trained, it can be extremely powerful. Deep learning is often used for things like speech recognition, understanding language, and that's what we're going to talk about next.

Natural Language Processing (NLP)

Hey, Vector, I have a question. What is natural language processing? Definition of natural language processing: a branch of information science that deals with natural language information. That's technically true, but it's one of those times where the dictionary definition really doesn't help very much. Natural language processing or NLP, is that area of artificial intelligence that deals with recognizing, understanding, analyzing, and even emulating the typical ways that humans communicate using either voice or text or both. Now these days, you don't just see the NLP term, but you'll see NLU for a natural language understanding, and NLG for natural language generation, and really these are just refinements. They make it clear what direction is a particular communication going. So if I'm speaking to or writing to the computer, it needs to understand my natural language, and if the computer is constructing some voice or text to send to me, then it needs to generate that natural language. So both NLU and NLG are part of NLP, and okay, a well-known example is using any of the various personal voice assistants. I don't want to wake up any device you may have, so I'll just say voice assistants like this or this or this or this, and when you talk to these devices, it's far more than just speech to text. It's more than just this basic dictation idea of, can the computer recognize a stream of individual words? No, the computer needs to actually understand what it is I'm talking about, and the variety of different ways I might phrase the same thought. Now when you were in school, you probably went through this process of taking a sentence and identifying the different parts; these are nouns, these are verbs, these are adverbs, and so on, but the goal of natural language understanding is

deeper. It's about identifying the meaning, the semantic content of the sentence, even when it's not a well-formed sentence at all. Was it a command? Was it a question? What's the subject? What are the keywords? And what is even the tone, the sentiment, or the emotion of this? That's very easy to think, well, yeah, but my company isn't making a new voice assistant so this doesn't apply to me, but it probably does, because the ability for AI to understand and generate natural language is becoming increasingly common in basic customer support situations, and in things like marketing and social media management. One example of NLU is in automated sentiment analysis. In a situation where you have online reviews or comments, to have a process that could scan and identify whether they're positive or neutral or negative, even perhaps making judgments about emotion; are any of these customers happy, are any of these customers angry, and deal with any major issues before they get out of hand. And beyond that, we're seeing more and more use of conversational interfaces, what are often called chatbots. The chatbots can be text based, and they can be embedded on a web page. They could be contained in a social network messaging app like Facebook, or you could perhaps interact with them by sending SMS text messages. And while early chatbots were very formulae and conscripted, very frustrating if you stepped outside the boundaries of exactly what they expected, these are becoming more and more capable of dealing with many of the basic, repeated, and predictable customer interactions. It's also becoming very common to integrate our own applications with voice assistants, like this or this. Think of how the expectations have changed over the last few years that maybe a decade ago an application might just be on the desktop, and then everybody expected to also have a web-based interface to it, and then also have a mobile interface to interact with it. And now we're getting to the point where people want conversational interfaces to that same app, but thankfully, if we want to do this, we don't have to build all this technology ourselves, because just as there are general machine learning frameworks, we can also find multiple options to build or integrate conversational user interfaces. Amazon AWS Office Cloud Services like Lex. There's also Polly, which can generate realistic-sounding speech. An IOS developer could use SiriKit to integrate their app with Apple's voice assistant. Microsoft Azure has the Language Understanding Service. Alibaba has Intelligent Speech Interaction; IBM Office, Watson natural language processing, and you can expect any AI platform will offer some kind of a specialized service about this. Let's actually use this to step a little more into the marketplace of AI solutions and AI vendors.

The AI Marketplace

I want to talk about the various vendors, products, and solutions on offer. But do keep in mind, this is an incredibly volatile industry and a volatile time in that industry. New startups appear all

the time. Companies get acquired. Names change. Products change. But what I can talk about is more generally how to look at this and how to think about what's on offer. But there is a category of AI-related companies I am going to intentionally avoid talking about, domain-specific areas like AI solutions targeted at healthcare or AI for the legal industry or solutions for transportation or hospitality or human resources simply because we would be here all day. To even provide a high-level overview of AI in healthcare, I could easily spend an hour on just that and barely scratch the surface. So for what's currently going on in your specific industry with AI and machine learning and deep learning and NLP, I will leave that as an exercise for the viewer. I'm going to focus here on some of the more generally applicable options. Now I said something in the machine learning section which I'm going to say again, the main reason for the incredible growth in applied AI over the last few years, the biggest enabler of all of this has been the large commercial cloud service providers offering these services. Now the order of the top companies changes slightly, depending on where you are in the world. But we can say Microsoft Azure, Amazon Web Services, or AWS, Google Cloud Platform, Alibaba Cloud, IBM Cloud, Oracle Cloud. There are others, but we'll leave it at this. They all offer a variety of pay-as-you-go artificial intelligence services you can tap into. They're ready to use together with all the computing power you need to run them, but you don't need to build your own data center and you only pay for what you use. And this has been what's reduced the barrier to entry and made it also very simple to get started. Now, to be clear, you don't have to use any of the cloud-based providers. There are open-source machine learning platforms and frameworks like TensorFlow and PyTorch, software a developer could just freely download and install. Now when I'm talking to people with some programming background but no experience in AI or machine learning, I'm often asked, What's the most common programming language for this? Now, recommending a language is always a contentious task, but for anyone with some background looking to take the next step or even just wanting to experiment, my default suggestion is simple, Python. Yes, there are other languages popular in the AI community, and if you're already working in a particular environment or using a specific technology that favors another language, let's say R, which is another common programming language in AI, then, of course, use that. But, otherwise, Python is never a bad choice. But if someone's intending to take that next up, there is one more aspect, not just how to do this, but how to do this responsibly.

Issues, Risks and Ethics in AI

We began this course with a cliche, and we're going to end with another one. AI is incredibly powerful, but with great power comes great responsibility. And we've become very aware in

recent years of harmful consequences of AI, sometimes intentional, sometimes unintentional. And one of the unintentional side effects would be algorithmic bias. One well publicized situation was where machine learning was used to try and streamline a recruitment and hiring process to attempt to recognize good candidates more quickly. Sounds fine. And the model was trained on 10 years of actual resumes, real data about the qualities of successful candidates. Sounds great. But this was in computing, which is male dominated. So just on pure numbers alone, what that meant is most successful candidates over the last 10 years were men, and that unintentionally became part of what the machine learning model thought it was supposed to look for. What does a successful candidate look like? A man. So does this new candidate match that description? Nope. Then score them lower. Now, this was a situation where there was no human prejudice involved, just the nature of a male-dominated profession. But it's where the model can end up not just reflecting, but indeed reinforcing that disparity. Figuring out when this is happening can actually be very difficult, because one of the very real issues is is that with many algorithms in machine learning, when we train a model, what you get from it is an answer. What you don't get is an explanation of that answer. This is sometimes called the black box of machine learning. You train the model with some existing data. You can then feed new data into the model, and it might say Candidate A is rated 2 out of 5. Candidate B is rated 5 out of 5. Why? It doesn't tell you that. You get the result; you don't get an explanation. And if you have an unquestioning culture of, well, that's what the computer says, then you might end up institutionalizing biases in the system without ever realizing it or ever having a chance to fix it. So we should always be looking at this as a tool. The machine learning is something to help us understand the data. It's not something that will understand it instead of us. We also have techniques like deepfakes, where AI can be used for intentional deception, disinformation. A deepfake usually refers to a video, but it can also be an image or even an audio recording. It's where AI can be used to either replace a person or make it seem like they said something they didn't. They generate extremely realistic results, and they're getting more realistic all the time. But because of these and other situations, there are already various initiatives around ethics in artificial intelligence and responsible AI. We have standards organizations like the IEEE that have their Global Initiative on the Ethics of Autonomous and Intelligence Systems. There's groups like Partnership on AI. This has over 100 member organizations, including Apple, Microsoft, Amazon, Google, Accenture, trying to develop guidelines and best practices about safety and transparency in artificial intelligence. Because there is an understanding that this is incredibly important. And it isn't just going to happen by itself. Because going back to that very first cliche that AI always ends up wanting to kill everybody, well, we've had people like Elon Musk, Bill Gates, and Stephen Hawking all raise serious concerns about what is likely to happen if we're not thinking about how to create this

ethical, responsible AI well in advance of it actually happening. Or as Professor Hawking said, "Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last." Interesting times ahead. Hope you enjoyed the course. See you next time. We hope you enjoyed this course. If you're interested in more content for technical leaders, content we keep short and focused with the up-to-date information you need to be informed and make decisions but without getting buried in the details, find other courses like this at plrsig.ht/exec.

Course author

Simon Allardice

Simon is a staff author at Pluralsight. With over three decades of software development experience, he's programmed in every discipline: from finance to transportation, nuclear reactors to game...

Course info

Level	Beginner
Rating	★★★★★ (36)
My rating	★★★★★
Duration	0h 40m
Released	11 Feb 2020

Share course