

If you don't have a technical background, but you still want to learn the basics of artificial intelligence, stick around because we were distilling Google's 4-hour AI course for beginners into just 10 minutes. I was initially very skeptical because I thought the course would be too conceptual. We're all about practical tips on this channel. And no, in Google, the course might just disappear after 1 hour. But I found the underlying concepts actually made me better at using tools like Chatchy BT and Google Bard, and cleared up a bunch of misconceptions I didn't know I had about AI, machine learning, and large language models. So, starting with the broadest possible question, what is artificial intelligence? It turns out, and I'm so embarrassed to admit I didn't know this. AI is an entire field of study, like physics. A machine learning is a subfield of AI. Much like how thermodynamics is a subfield of physics. Going down another level, deep learning is a subset of machine learning, and deep learning models can be further broken down into something called discriminative models and generative models. Large language models, LLMs, also fall under deep learning and right at the intersection between generative and LLMs, is the technology that powers the applications we're all familiar with. Chatchy BT and Google Bard. Let me know in the comments if this was news to you as well. Now that we have an understanding of the overall landscape and you see how the different disciplines sit in relation to each other, let's go over the key takeaways you should know for each level. In the nutshell, machine learning is a program that uses input data to train a model. That trained model can then make predictions based on data it has never seen before. For example, if you train a model based on Nike sales data, you can then use that model to predict how well a new shoe from Adidas would sell based on Adidas sales data. Two of the most common types of machine learning models are supervised and unsupervised learning models. The key difference between the two is supervised models use labeled data and unsupervised models use unlabeled data. In this supervised example, we have historical data points that plot the total bill amount at a restaurant against the tip amount. And here the data is labeled. Blue dot equals the order was picked up and yellow dot equals the order was delivered. Using a supervised learning model, we can now predict how much tip we can expect for the next order given the bill amount and whether it's picked up or delivered. For unsupervised learning models, we look at the raw data and see if it naturally falls into groups. In this example, we plotted the employee tenure at a company against their income. We see this group of employees have a relatively high income to years work ratio versus this group. We can also see all these are unlabeled data. If they were labeled, we would see male, female, years worked, company function, etc. We can now ask this unsupervised learning model to solve a problem like if a new employee joins, are they on the fast track or not? If they appear on the left, then yes. If they appear on the right, then no. Pro tip. Another big difference between the two models is that after a supervised learning model makes a prediction, it will compare that prediction to the training data used to train that model. And if there's a difference, it tries to close that gap. Unsupervised learning models do not do this. By the way, this video is not sponsored, but it is supported by those of you who subscribed to my paid productivity newsletter on Google tips. Link in the description if you want to learn more. Now we have a basic grasp of machine learning. It's a good time to talk about deep learning, which is just a type of machine learning that uses something called artificial neural networks. Don't worry. All you have to know for now is that artificial neural networks are inspired by the human brain and looks something like this. Layers of nodes and neurons. And the more layers they are, the more powerful the model. And because we have these neural networks, we can now do something called semi-supervised learning, whereby a deep learning model is trained on a small amount of label data and a large amount of unlabeled data. For example, a bank might use deep learning models to detect fraud. The bank spends a bit of time to tag or label 5% of transactions as either fraudulent or not fraudulent. And they leave the remaining 95% of transactions unlabeled because they don't have the time or resources to label every transaction. The magic happens when the deep learning model uses the 5% of label data to learn the basic concepts of the task. Okay, these transactions are good and these are bad. Okay. Apply those learnings to the remaining 95% of unlabeled data. And using this new aggregate data set. The model makes predictions for future transactions. That's pretty cool. And we're not done because deep learning can be divided into two types. Discriminative and generative models. Discriminative models learn from the relationship between labels of data points and only has the ability to classify

those data points. Fraud, not fraud. For example, you have a bunch of pictures or data points. You purposely label some of them as cats and some of them as dogs. A discriminative model will learn from the label, Catardog. And if you submit a picture of a dog, it will predict the label for that new data point. A dog. We finally get to generative AI. Unlike discriminative models, generative models learn about the patterns in the training data. Then after they receive some input, for example, a text prompt from us, they generate something new based on the patterns they just learned. Going back to the animal example, the pictures or data points are not labeled as Catardog, so a generative model will look for patterns. Oh, these data points all have two years, four legs, a tail, likes dog food, and barks. When asked to generate something called a dog, the generative model generates a completely new image based on the patterns it just learned. There's a super simple way to determine if something is generative AI or not. If the output is a number, a classification, spam, not spam, or a probability, it is not generative AI. It is gen AI when the output is natural language, text, or speech, an image, or audio. Basically, generative AI generates new samples that are similar to the data it was trained on. Moving on to different generative AI model types, most of us are familiar with text-to-text models like Chagy B.T. and Google Bard. Other common model types include text-to-image models like mid-journey, dolly, and stable diffusion. These can not only generate images, but edit images as well. Text-to-video models, surprise surprise, can generate and edit video footage. Examples include Google's image and video, cog video, and the very creatively named Make a Video. Text-to-3D models are used to create game assets, and a little known example would be OpenAI's shape-e model. And finally, text-to-task models are trained to perform a specific task. For example, if you type add Gmail, summarize my unread emails, Google Bard will look through your inbox, and summarize your unread emails. Moving over to large language models, don't forget that LLMs are also a subset of deep learning, and although there is some overlap, LLMs and Gen AI are not the same thing. An important distinction is that large language models are generally pre-trained with a very large set of data, and then fine-tune for specific purposes. What does that mean? Imagine you have a pet dog. It can be pre-trained with basic commands like sit, come, down, and stay. It's a good boy and a generalist, but if that same good boy goes on to become a police dog, a guide dog, or a hunting dog, they need to receive specific training, so they're fine-tune for that specialist role. A similar idea applies to large language models. They're first pre-trained to solve common language problems like text classification, question answering, document summarization, and text generation. Then, using smaller industry-specific datasets, these LLMs are fine-tuned to solve specific problems in retail, finance, healthcare, entertainment, and other fields. In the real world, this might mean a hospital uses a pre-trained large language model from one of the big tech companies, and fine-tunes that model with its own first-party medical data, to improve diagnostic accuracy from X-rays and other medical tests. This is a win-win scenario because large companies can spend billions developing general-purpose large language models, then sell those LLMs to smaller institutions like retail companies, banks, hospitals, who don't have the resources to develop their own large language models, but they have the domain-specific datasets to fine-tune those models. Trote tip, if you do end up taking the full course, I'll link it down below, it's completely free. When you're taking notes, you can right-click on the video player and copy video URL at the current time, so can quickly navigate back to that specific part of the video. There are five modules total and you get a badge after completing each module. The content overall is a bit more than the theoretical side, so you definitely want to check out this video on how to master prompting next. See you on the next video then mean time. Have a great one.