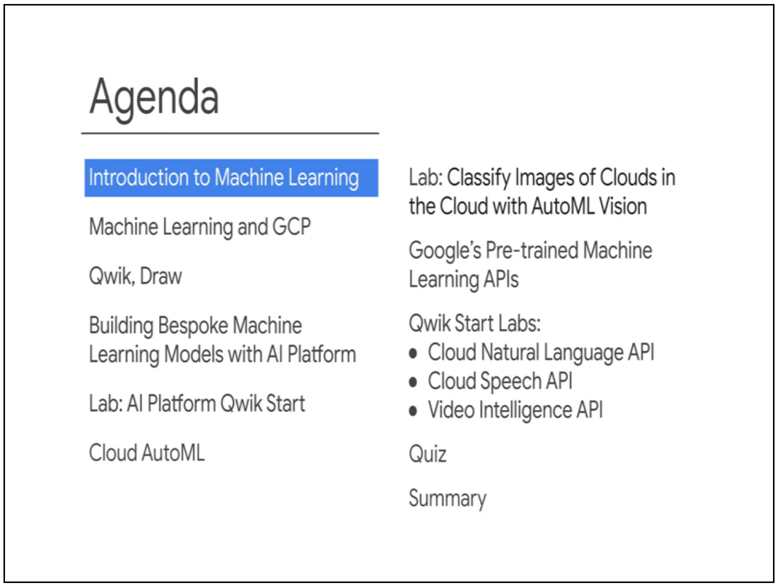
Google Cloud Computing Foundation Course
Evan Jones
Technical Curriculum Developer
Google Cloud

Lecture-77
Introduction to ML

(Refer Slide Time: 00:07)



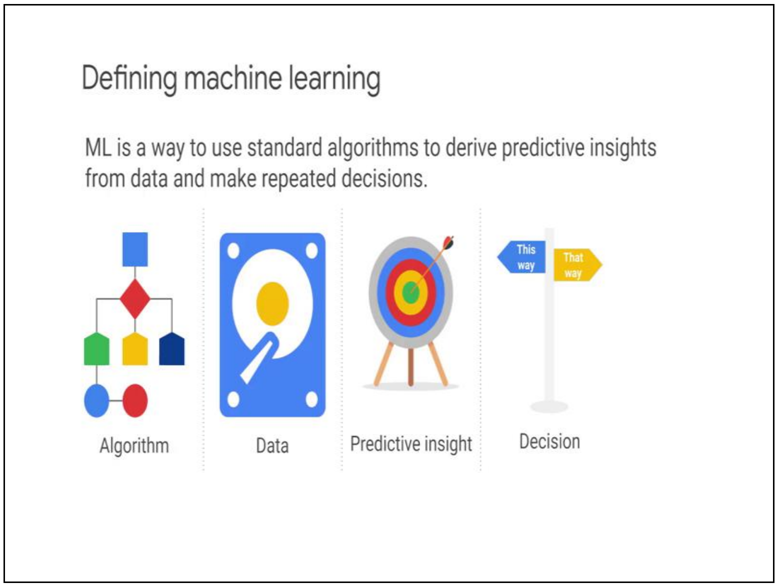
The world is filled with things that were able to react to, and understand with how much thought. For example, consider a stop sign that partially covered by snow. It is still a stop sign or a chair that is five times bigger than usual and it is still a place to sit. But for computers who do not have the benefit of growing up and learning the nuances of these objects the world is often and much more messy and complicated.

For the first topic, to start with this video, making sense of a messy world or Google engineers and researchers discussed how machine learning is beginning to make computers and many of the things that we use in for such as Maps, search recommending videos, translations and so on much better.

(Video Start Time: 00:47)

(Video End Time: 04:32)

(Refer Slide Time: 04:34)



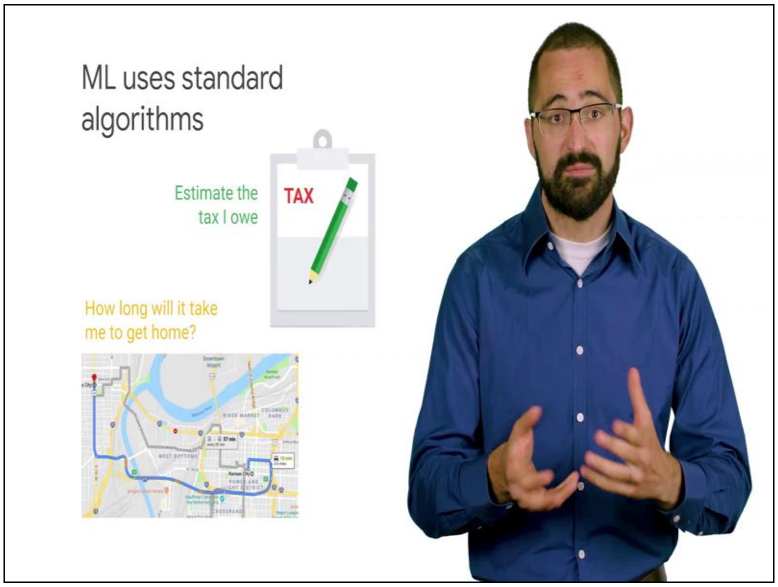
Steve heard a lot about machine learning or ML. Let us start with a definition. What is ML definition? Here is a definition I like to use. ML is a way to get predictive insights from data to make repeated decisions off. You do this, using algorithms that are relatively general and applicable to a wide variety of data sets. Pick up a typical company and how they use their data today. Perhaps they have a dashboard that business analysts and decision-makers view on a daily basis or report that is read on a monthly basis. This is an example of a backward-looking use of data.

Looking at historical data, to create reports and dashboards, this is what people tend to mean when they talk about BI or business intelligence. A lot of data analytics is backward-looking nothing wrong with that. Instead we use ML or Machine Learning to generate forward-looking or predictive insights. Of course, the point of looking at historical data might be to make those decisions. Perhaps, business analyst examines the data and they suggest new policies or rules.

Just for example, that is possible to raise the price of a product in a certain region now that business and he is making unpredicted insight, but is that scalable? Can a business analyst make such a decision for every single product in every single region? And can they be dynamically adjusts the price every second? Ah, here is where the computers get involved. In order to make decisions around predictive insights repeatedly, you need ML.

I need a computer program to derive those insights for you. So ML is about making predictive insights from data, many of the main time. It is about scaling up BI and decision-making. The other part of the machine on definition is around the use of standard algorithms. Emily uses standard algorithms to solve when look like seemingly different problems. Normally when we think of computers, we think a program that do different things.

(Refer Slide Time: 06:39)



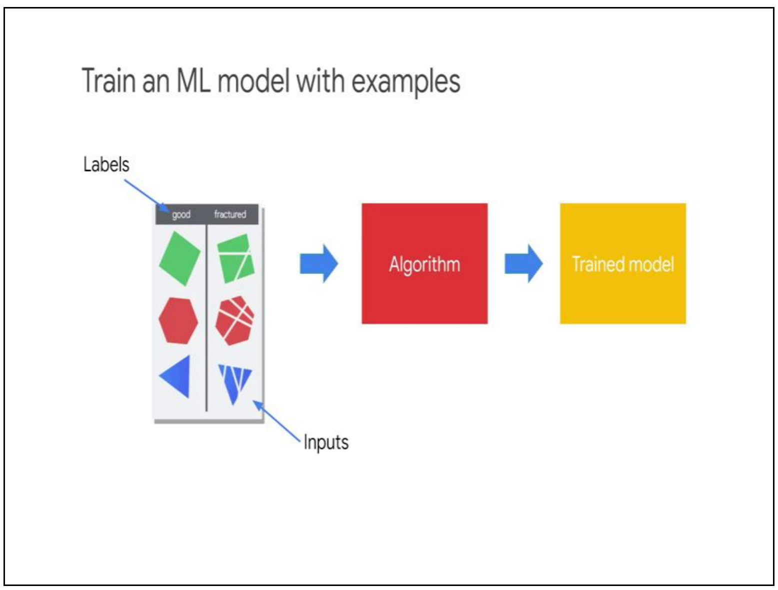
For example, it is often used to file your taxes is very different from the software that used to get directions home when you are driving. Machine learning is a little different. You use the same software under the hood. That is what we mean when we say ML uses standard algorithms, but you can train that software to do very different things. You can train the software to estimate the amount of taxes that you owe. Or train that to offer to estimate the amount of time it will take to get you home.

This ML software once trained on your specific use case is called a model. So, you know the model that can estimate your taxes; a model they can estimate the time to get you home. We use the term model because it is an approximation. It is a model of reality. For example, we are giving the computer lots of historical data on drive time to New York City and it will earn their relationships in the data traffic patterns, seasonality, Time of day impact to predict today's commute time home.

(Refer Slide Time: 07:41)

Whatever the domain ML modeling requires lots of training examples. We will train the model to estimate tax by showing it many, many, many examples of Prior year tax returns. Or train the model to estimate trip duration by showing it many, many, many different Journeys. So, the first stage of ML is to train ML model with lots of good examples.

(Refer Slide Time: 08:08)

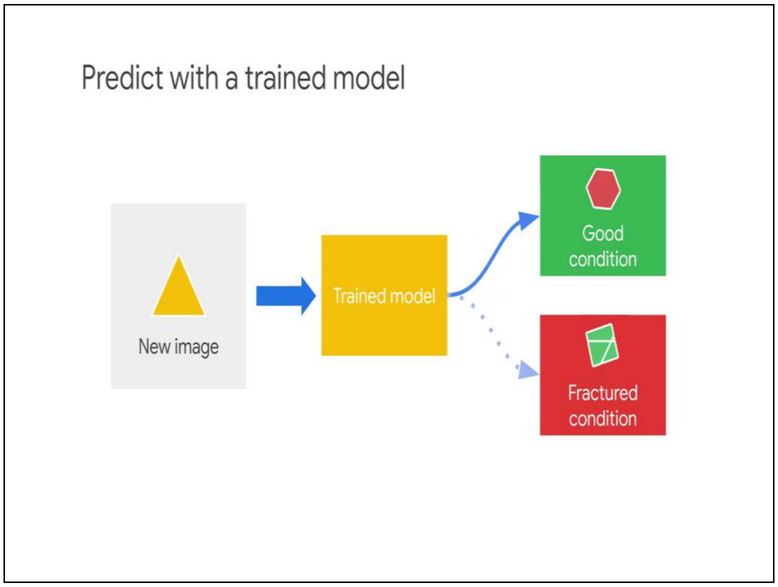


An example consists of an input in the correct answer for the input. That is called the label. In the case of structured data that has rows and Columns of data and input can simply be a single row of data. In unstructured data, like images, an input, be a single image, say like a cloud that you want to classify: Is this a rain cloud or is this not? Now, imagine you work for a

Manufacturing Company. You want to train a machine learning model to detect defects in these parts, before they get assembled into the final products for users.

Now you can start by collecting a dataset of the images for these parts. Some of the parts to be good, some of these parts can be fractured or broken up. And for each image you assign the corresponding label. That is the right answer broken or not broken part and then use this set of examples as training data for your model.

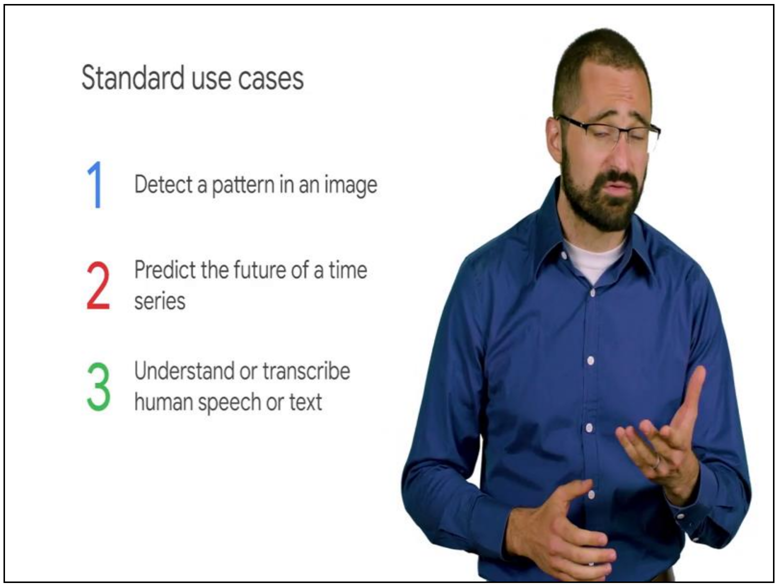
(Refer Slide Time: 09:06)



After you train the model, you can then use it to predict the label of images that it has never seen before. Learn from the past, predict for the future. Here, your input for the train model is an image of the park because the model has already been trained is correctly able to predict at this part is in good condition know that the image here is different from the ones used in our training examples, but it still works because the ML model has generalized.

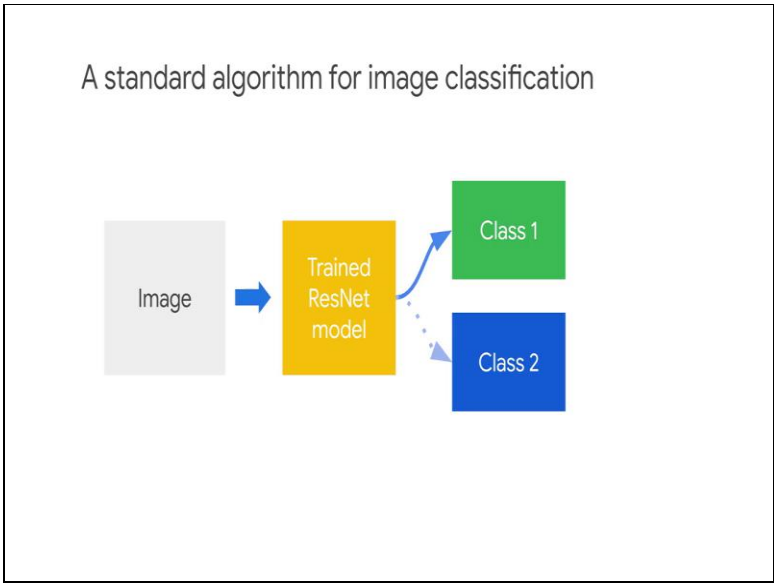
It has not memorized the training data those specific examples that you showed him and is learned a more general idea of what a good-looking part with a good condition for that part looks like. So, why do we say these algorithms are standard? Algorithms exist independently of your case even though it detecting manufacturing defects in Parts in terms of images and detecting something like disease, leaves and tree images or two very different use cases.

(Refer Slide Time: 10:03)



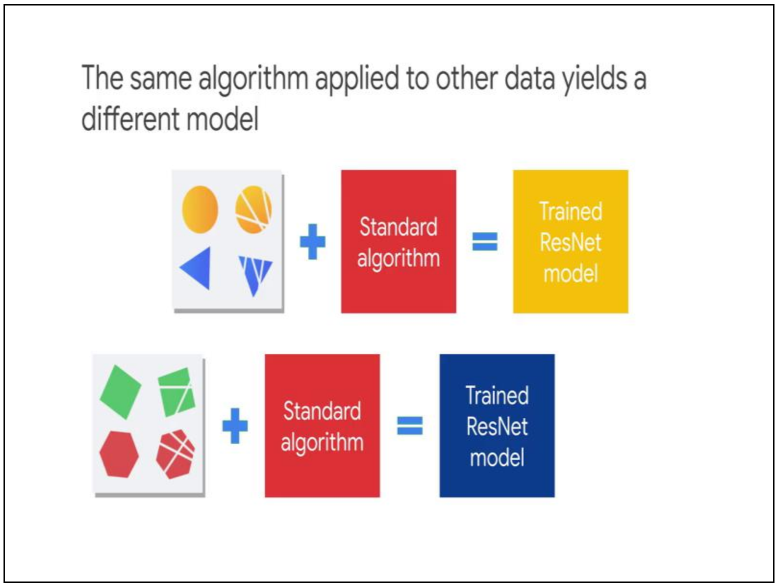
The same algorithm, an image classification network works for both. Similarly, there are standard algorithms for predicting the future value of a Time series data set, how to transcribe human speech to text.

(Refer Slide Time: 10:21)



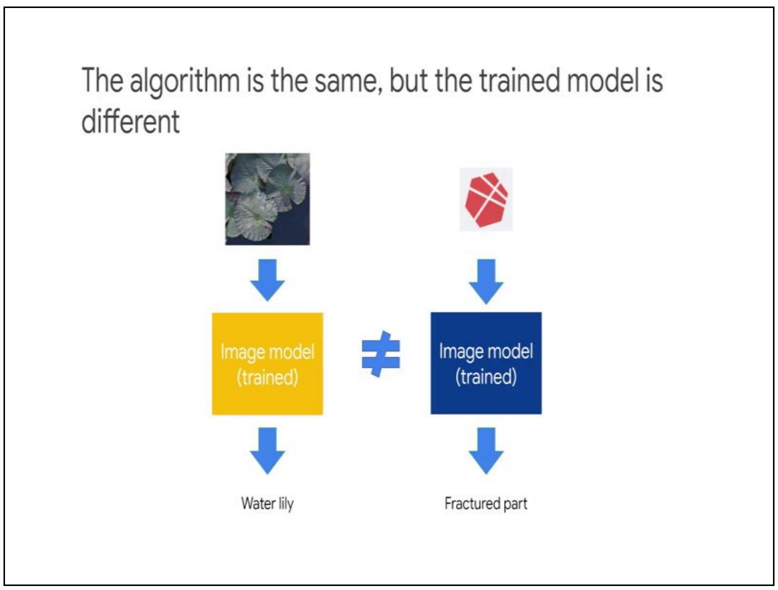
ResNet is a standard algorithm for image classification. There is not crucial to understand how an image classification algorithm works. Only that is the hour that you should use, if you need to classify images of Automotive Parts. When you use the same algorithm on different data sets,

(Refer Slide Time: 10:37)



There are different features or input relative to the different used cases and you can see them represented visually here.

(Refer Slide Time: 10:48)



You might be asking yourself, Is it not the logic different? You cannot possibly use the same rules for identifying defects and Manufacturing that you do and identifying different types of leaves. You are right. The logic is different, but ML does not use logical if-then rules. The image classification network is not like that set of rules if this then that but a function that learn how to distinguish between categories of images.

You know, we start with the same standard algorithm after training the trained model that classifies leaves is different from the trained model that classifies manufacturing Parts. Guess what, you can actually reuse the same code for the other use cases focused on the same kind of task. So in our example, we are identifying manufacturing defects with a higher level tax with classifying images.

You can reuse the same code for another image classification problem, like finding examples of your products in photos posted on social media. However, you still at the train it separately for each use case.

(Refer Slide Time: 11:56)

The main thing to know is that for machine learning, your model will only be as good as your data. And more often than or not, you use a lot of data for machine learning. For example, that we talked about you will need a large dataset of historical examples of both rejected parts and parts in good condition in order to train a model to categorize the parts are defective or not.

The basic reason why ML models need high-quality training data is because they do not have human general knowledge like we do. Data is the only thing that they have access to, to learn from.