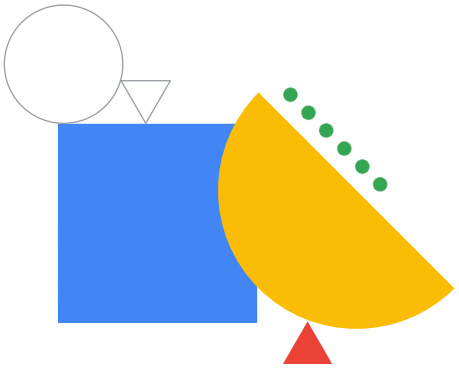
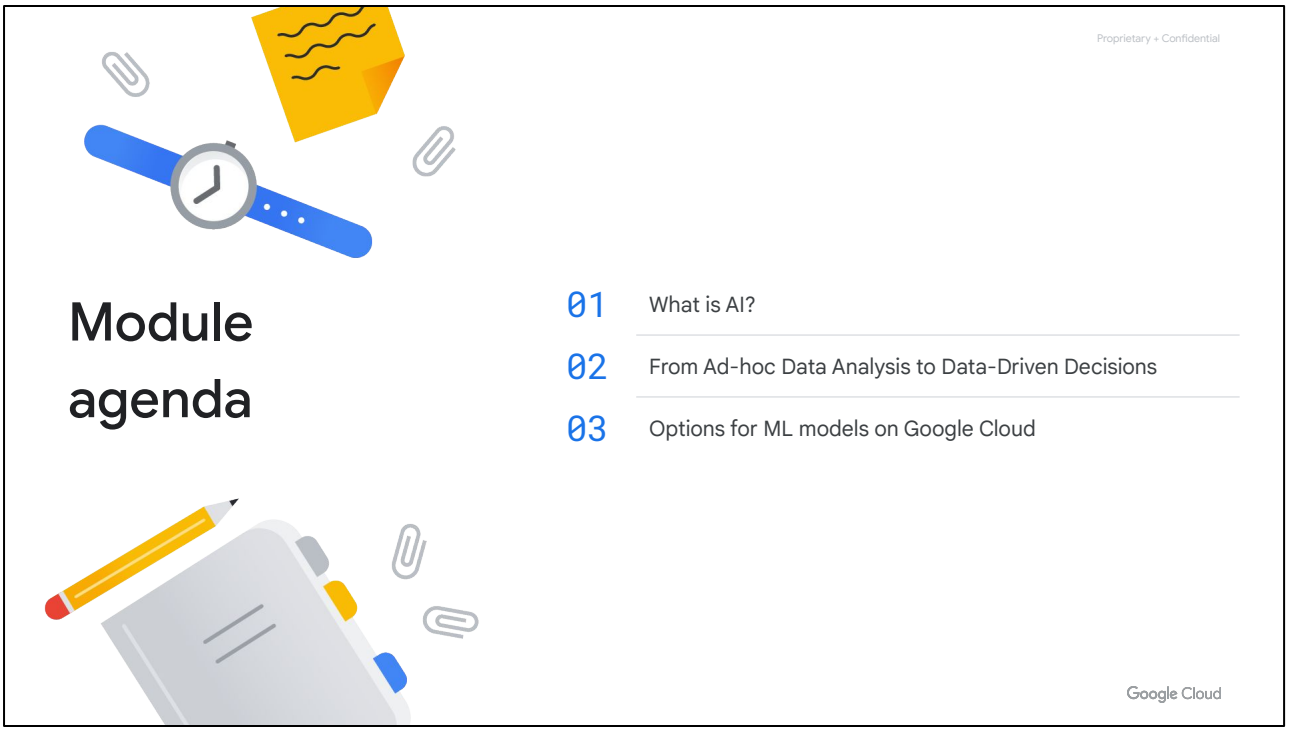
Google Cloud

Introduction to Analytics and AI



In this first module, we'll explore what AI is and its role in supporting data-driven analytics.

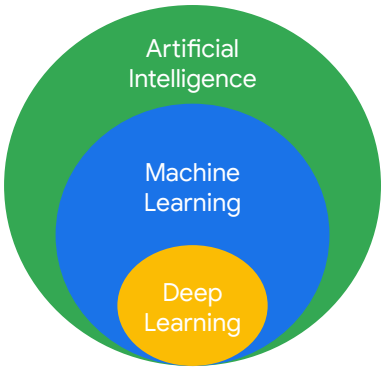


We'll start by answering the question: What is AI? Then, we'll discuss some of the process of moving from ad-hoc analysis to data-driven decision making. Finally, we will cover some of the options that exist to leverage machine learning on Google Cloud.

What is AI?

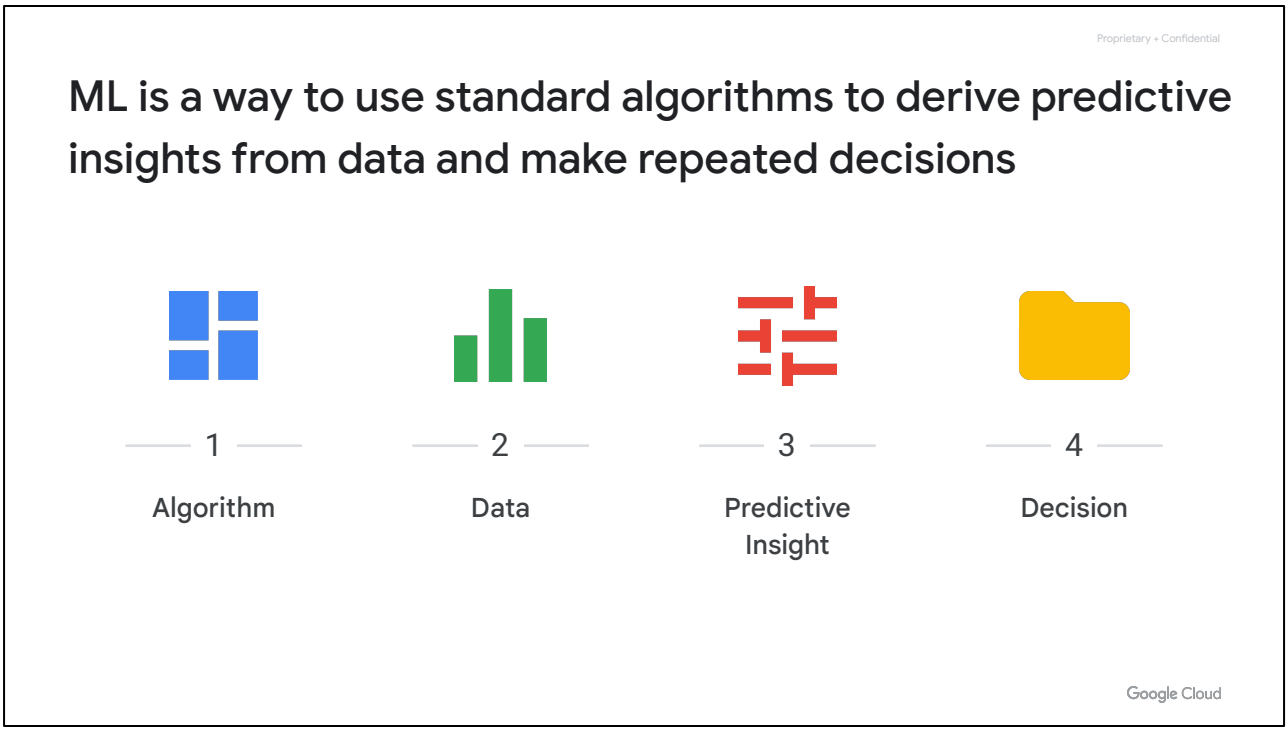
So let’s start by examining what AI is.

Machine Learning is a type of AI, and deep learning is a type of machine learning



Class of problems we can solve when computers think/act like humans

A very common question asked is: What's the difference between Artificial Intelligence - AI, Machine Learning - ML and Deep Learning? Well, one way to think about it is AI is that it’s a discipline - something like physics. AI refers to machines that are capable of acting autonomously, machines that think. AI has to do with the theory and methods to build machines that can solve problems by thinking and acting like humans.



You've likely heard a lot about Machine Learning or ML. Let's start with a definition. What is ML?

1. ML is a way to use standard **algorithms** to derive predictive insights from data and make repeated decisions. You do this by applying algorithms that are relatively general and applicable to a wide variety of datasets.

2. Think of 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 a report that's 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.

3. A lot of data analytics is backward-looking, nothing wrong with that. But instead we're going to use ML or Machine Learning to generate forward-looking or **predictive insights**.

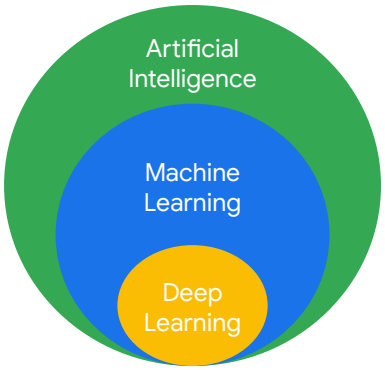
4. Of course, the point of looking at historical data might be to make those **decisions**. Perhaps business analysts examine the data and they suggest new policies or rules. They could suggest for example that it’s possible to raise the price of a product in a certain region.

Now, that business analyst is making a predictive insight but is that scalable? Can the

business analyst make such a decision for every single product in every single region? And can they dynamically adjust the price every second?

Here's where the computers get involved. In order to make decisions around predictive insights repeatable, you need ML. You need a computer program to derive those insights for you. So ML is about making predictive insights from data, many of them at a time. It's about scaling up BI and decision-making.

Why are Machine Learning and Deep Learning so exciting?



Class of problems we can solve when **computers think/act like humans**

Scalably solve those problems using **data** examples **(not custom code)**

Even when that data consists of **unstructured data** like images, speech, video, natural language text, etc.

Machine Learning is a tool set like Newton's Laws of Mechanics. Just as you can use Newton's laws to figure out how long it'll take a ball to drop when it falls off a cliff, you can use Machine Learning to scalably solve certain kinds of problems using data examples. But without the need for any custom code.

Deep Learning is a type of Machine Learning that works even when the data consists of unstructured data, like images, speech, video, natural language text and so on. One kind of deep learning is image classification. A machine can learn how to classify images into categories when it's shown lots of different examples. And one really powerful thing about deep learning is that oftentimes, in really complex problems, it can do better than a human. The basic difference between Machine Learning and other techniques in AI is that in Machine Learning machines learn. They don't start out intelligent, they become intelligent.

Keller Williams uses AutoML Vision to automatically recognize common elements of house furnishings and architecture

AutoML Vision

Let's take a look at a real-world example. Keller Williams, a U.S. real estate company, uses AutoML Vision to automatically recognize specific features of houses like built-in bookcases.

This helps agents get houses listed faster, and buyers find houses that meet their needs.

Neil Dholakia, Chief Product Officer stated that "By training a custom model to recognize common elements of furnishings and architecture, customers can automatically search home listing photos for specific features like granite countertops which are more modern."

This application of machine learning quickly allows Keller Williams realtors to record a video walkthrough of a new home and use the object detection capabilities of AutoML Vision to find and tag key aspects of the home that customers would want to search on.

A big benefit for their organization is that they already had many existing images and videos of home walkthroughs already. They simply fed them into the pre-built AutoML Vision model and customized it. All without writing a line of code. You'll learn more about AutoML Vision and practice creating models with it later in this course.

[https://cloud.google.com/blog/products/gcp/empowering-businesses-and-developers-do-more-ai]

Kewpie uses ML to sort out the bad potatoes in baby food

kewpie

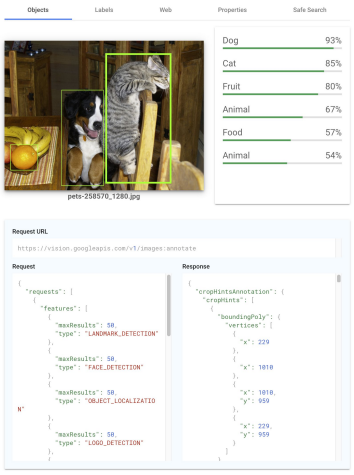
Original process required humans to identify low-quality ingredients, which was expensive and stressful.

Machine learning was used to replicate the quality control process.

Japanese manufacturer Kewpie produces baby food among other products. Kewpie found that quality is not necessarily a matter of safety—because the food itself is safe—but discoloration can concern parents. So Kewpie turned to Google and our partner Brainpad to build a solution that leverages image recognition to detect low-quality potato cubes. The ML algorithm enabled Kewpie to free people from the tiring work of inspecting potatoes and focus on other higher value tasks.

[https://www.blog.google/products/google-cloud/how-ai-can-help-make-safer-baby-food-and-other-products/]

Play around with the power of AI yourself...

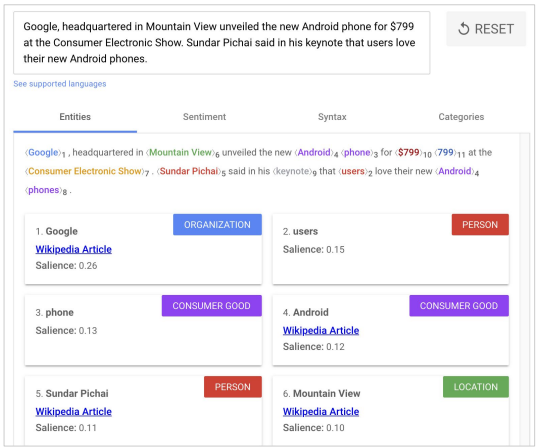


cloud.google.com/vision/

- Object detection
- Labeling and confidence
- Web lookup
- Pre-trained (call the API)

You can easily explore the power of AI yourself. Navigate to cloud.google.com/vision and try some tests in your browser with the Vision API by uploading an image from your device.

Try Google's natural language API



cloud.google.com/natural-language/

- Entity extraction
- Sentiment analysis
- Sentence structure
- Pre-trained (call the API)

You can also test out Google's Natural Language API in your browser by visiting cloud.google.com/natural-language

From Ad-hoc Data Analysis to Data-Driven Decisions

In this lesson, we'll explore how AI can help us to move from ad-hoc data analysis to data-driven decisions.

Imagine you're the owner of a bicycle rental business (in London). How do you stock enough bicycles?

Commuter Bikes

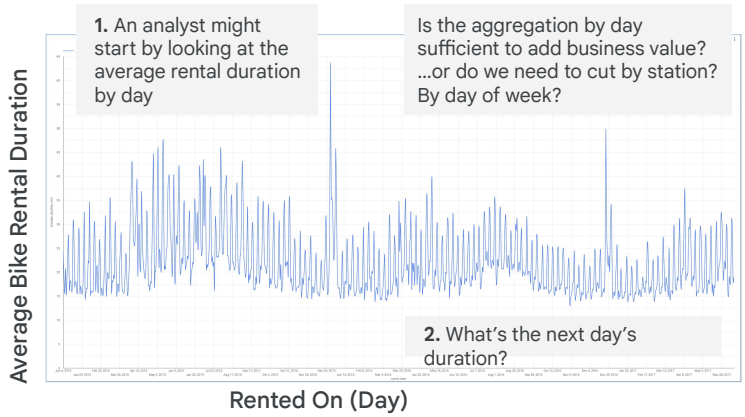
If rental is likely to be for a **short duration**, we need to have commuter bikes in stock

Road Bikes

If rental is likely to be for a **long duration**, we need to have road bikes in stock

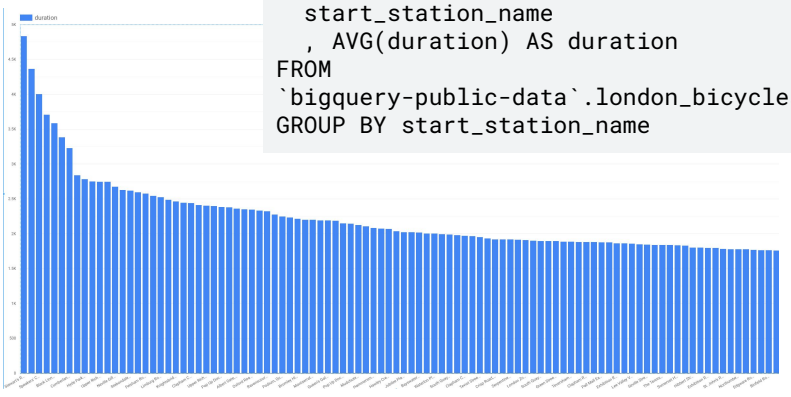
To help illustrate the journey from ad-hoc to data-driven decisions, consider the scenario of a bicycle rental business owner in London. How can you ensure you have enough but not too many bicycles on hand?

You hire a data analyst to help get you insights on how to keep the right bicycles in stock



You can use data analysis to look at trends in your historical data. This could include exploring the average daily rental duration over time.

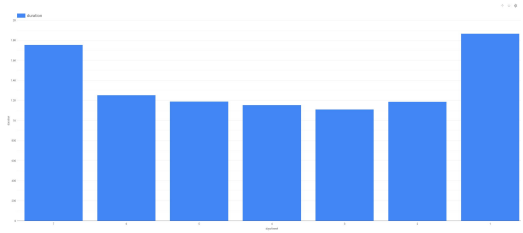
Does the duration of a rental vary by station?



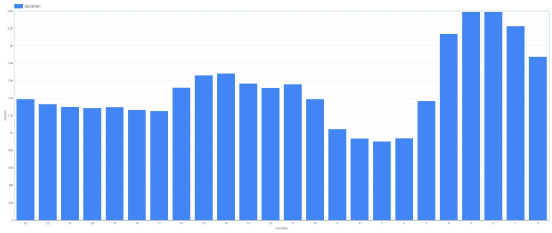

```
SELECT
  start_station_name
  , AVG(duration) AS duration
FROM
`bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY start_station_name
```


In addition, the analysis may consider the extent to which rental durations vary by station.

How about the day of the week? Hour of day?



```
SELECT
  EXTRACT(dayofweek
  FROM
    start_date) AS dayofweek,
  AVG(duration) AS duration
FROM
 `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
  dayofweek
```



```
SELECT
  EXTRACT(hour
  FROM
    start_date) AS hourofday,
  AVG(duration) AS duration
FROM
 `bigquery-public-data`.london_bicycles.cycle_hire
GROUP BY
  hourofday
```


You can even drill down into very specific ranges of time, such as the day of the week, or hour of any given day.

This ad-hoc analysis is great but...

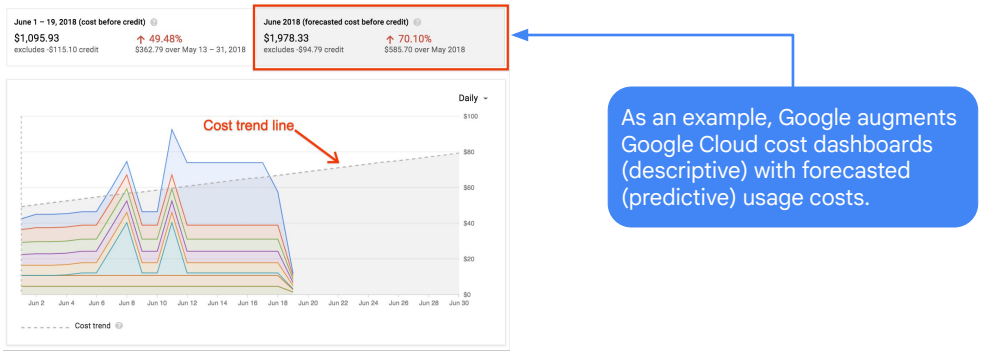
- A lot of manual, repetitive work involved for the data analyst
- Any decisions made will be based on hunches on how all these factors interact
- Wouldn't it be better if we could automate this analysis?

But it requires a lot of work sifting through the data for perhaps no meaningful insights.

Wouldn't it be easier to automate this type of work?

… what we need is an ML model to be able to make predictions

Goal: Augment our dashboards with predicted values e.g. prediction for the duration of a rental



The answer is using an ML model to make predictions about future needs based on prior usage.

[https://cloudplatform.googleblog.com/2018/07/predict-your-future-costs-with-google-cloud-billing-cost-forecast.html]

Use the ML model to anticipate what type of bike/how many to stock at your locations

- The ML model takes some of the drudgery out of ad-hoc analysis to help you make truer data-driven decisions.
- Can build a ML model in BigQuery or Vertex AI.

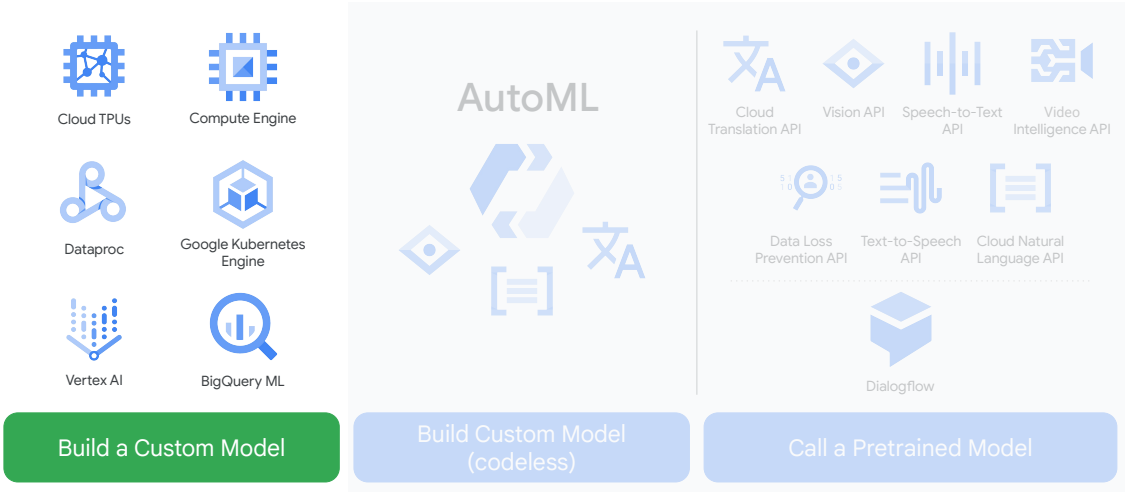
```
CREATE OR REPLACE MODEL
  bike_model.model_bucketized TRANSFORM(* EXCEPT(start_date),
  IF
    (EXTRACT(dayofweek
      FROM
        start_date) BETWEEN 2 AND 6,
      'weekday',
      'weekend') AS dayofweek,
    ML.BUCKETIZE(EXTRACT(HOUR
      FROM
        start_date),
      [5, 10, 17]) AS hourofday )
OPTIONS
  (input_label_cols=['duration'],
    model_type='linear_reg') AS
SELECT
  duration,
  start_station_name,
  start_date
FROM
  `bigquery-public-data`.london_bicycles.cycle_hire
```


An ML Model can apply different algorithms to derive insights on a much larger scale and faster than by human intervention. The example shown uses a linear regression model against hourly sections of bike rental data.

Options for ML Models on Google Cloud

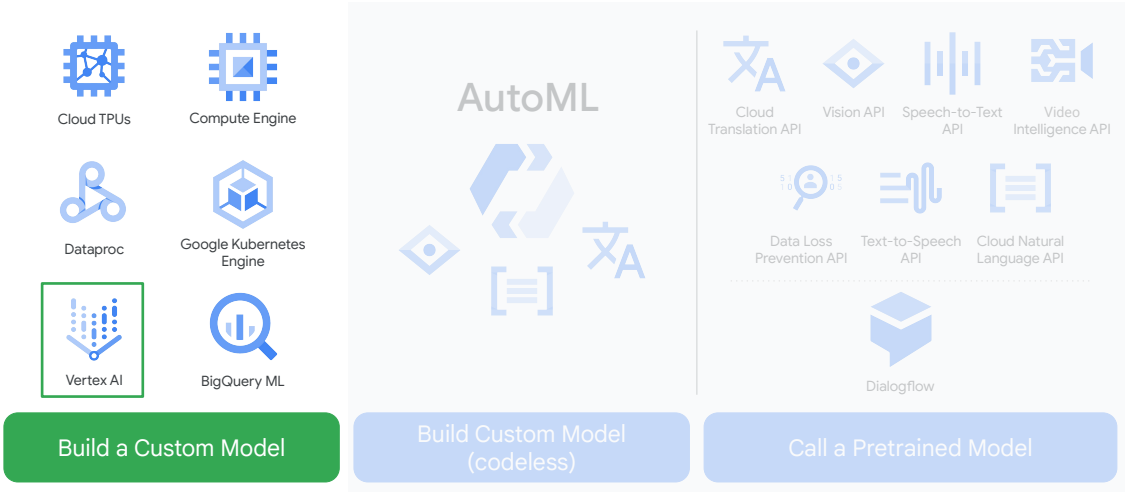
Finally, let's identify some of the options that exist to leverage machine learning on Google Cloud.

Leverage pre-trained models or build your own

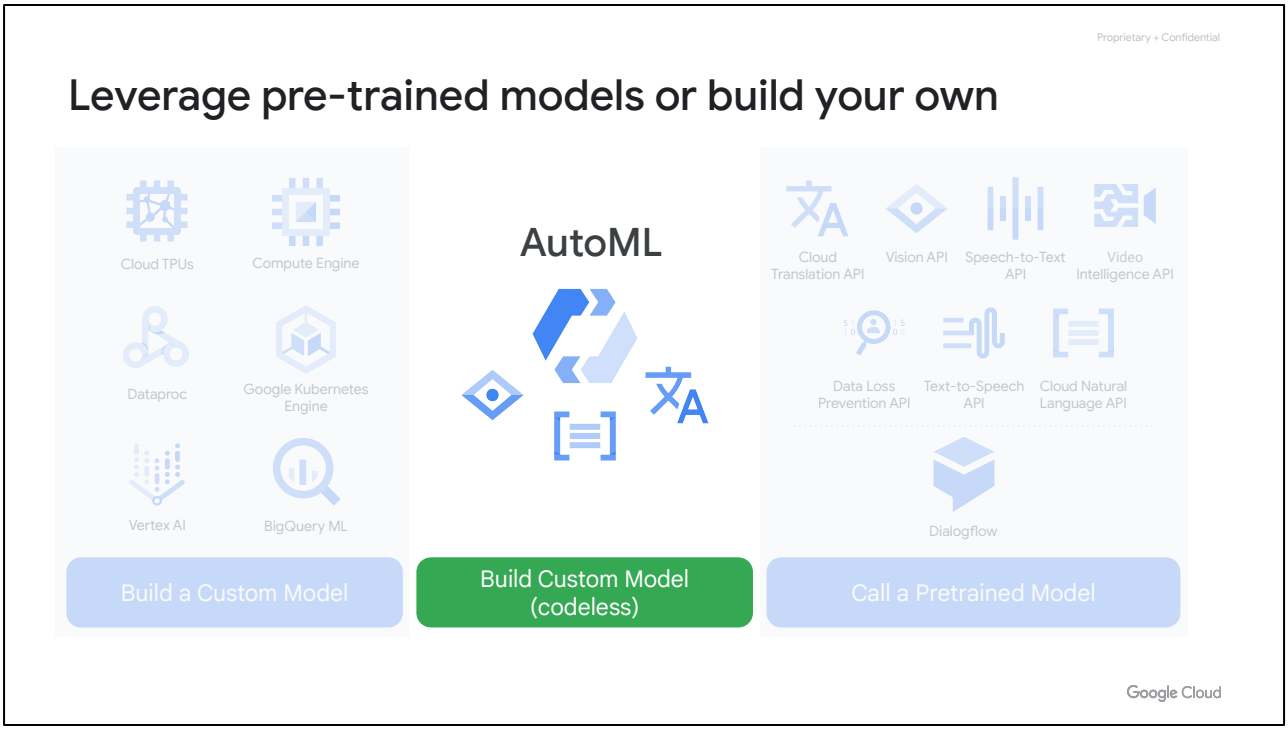


Advanced users who want more control over the building and training of their ML models will use tools that offer the levels of flexibility that they're looking for.

Leverage pre-trained models or build your own

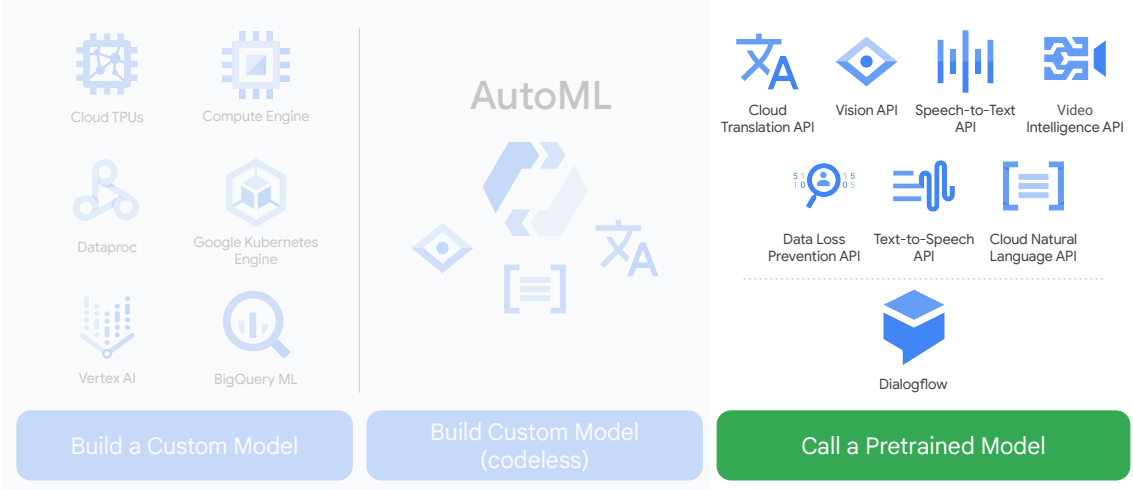


This can involve developing custom models through an ML library, like TensorFlow, that is supported on Vertex AI. This option works well for data scientists with the skills and the need to create a custom model.



But increasingly, you don't have to do that. Google makes the power of machine learning available to you even if you have a limited knowledge of machine learning. You can use AutoML to build on Google's machine learning capabilities, to create your own custom machine learning models that are tailored to your specific business needs. And then, integrate those models into applications and websites. All without running a line of TensorFlow code.

Leverage pre-trained models or build your own



Alternatively, Google has a range of pre-trained, meaning you don't need to bring your own data, machine learning models that are ready for immediate use within applications in ways that the respective APIs are designed to support. Such pre-trained models are excellent ways to replace user input with machine learning.