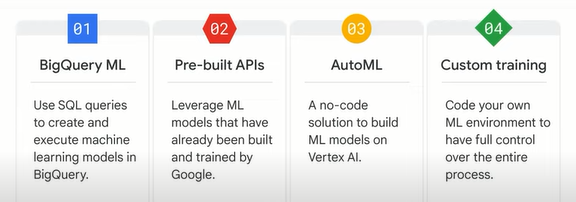
**Machine Learning Option on Google Cloud.**

**Introduction.**

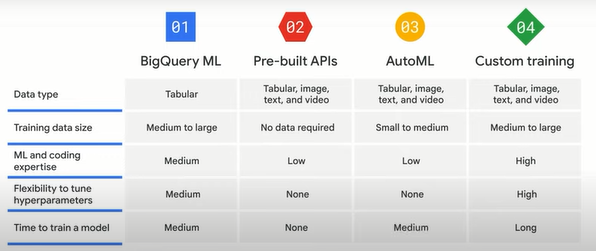
in previous sections of this course you learned about many data engineering tools available from google cloud now let's switch our focus to machine learning in this section we'll explore the different options google cloud offers for building machine learning models additionally we'll explain how a product called vertex ai can help solve machine learning challenges so you might be wondering why should i trust google for artificial intelligence and machine learning google is an ai first company and is recognized as a leader across industries because of its contributions in the fields of artificial intelligence and machine learning in 2021 google was recognized as a leader in the gartner magic quadrant for cloud ai developer services and in recent years has also received recognition in numerous annual industry awards and reports and at google we've been implementing artificial intelligence for over 10 years into many of our critical products systems and services for example have you ever noticed how gmail automatically suggests three responses to a received message this feature is called smart reply which uses artificial intelligence to predict how you might respond behind this intelligence is ai technology known as natural language processing which is just one example of an impressive list of technologies that google scientists and engineers are working on we'll explore these in more depth later in the course the goal of these technologies is not for exclusive use to only benefit google customers the goal is to enable every company to be an ai company by reducing the challenges of ai model creation to only the steps that require human judgment or creativity so for workers in the travel and hospitality field this might mean using ai and ml to improve aircraft scheduling or provide customers with dynamic pricing options for retail sector employees it might mean using ainml to leverage predictive inventory planning the potential solutions are endless what are the problems in your business that artificial intelligence and machine learning might help you solve take a moment to think about this question before continuing to the next video.

**Options to build ML models.**

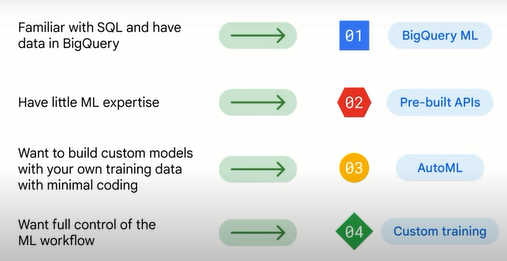
google cloud offers four options for building machine learning models the first option is bigquery ml you'll remember from an earlier section of this course that bigquery ml is a tool for using sql queries to create and execute machine learning models in bigquery if you already have your data in bigquery and your problems fit the predefined ml models this could be your choice the second option is to use pre-built apis which are application programming interfaces this option lets you leverage machine learning models that have already been built and trained by google so you don't have to build your own machine learning models if you don't have enough training data or sufficient machine learning expertise in-house the third option is automl which is a no code solution so you can build your own machine learning models on vertex ai through a point-and-click interface and finally there's custom training through which you can code your very own machine learning environment the training and the deployment which gives you flexibility and provides full control over the entire process



let's compare the four options to help you decide which one to use for building your ml model data type bigquery ml only supports tabular data while the other three support tabular image text and video data training data size pre-built apis don't require any training data while bigquery ml and custom training require a large amount of data machine learning and coding expertise pre-built apis and automl are user-friendly with low requirements while custom training has the highest requirement and bigquery ml requires you to understand sql flexibility to tune the hyper parameters at the moment you can't tune the hyper parameters with pre-built apis or automl however you can experiment with hyper parameters using bigquery ml and custom training time to train the model pre-built apis require no time to train a model because they directly use pre-built models from google the time to train a model for the other three options depends on the specific project normally custom training takes the longest time because it builds the ml model from scratch unlike automl and bigqueryml



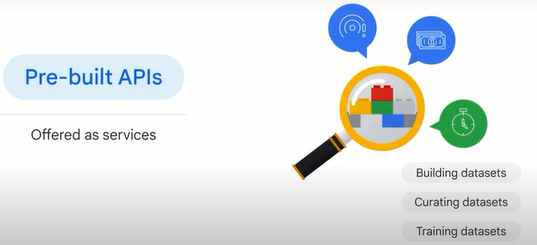
selecting the best option will depend on your business needs and ml expertise if your data engineers data scientists and data analysts are familiar with sql and already have your data in bigquery bigquery ml lets you develop sql based models if your business users or developers have little ml experience using pre-built apis is likely the best choice pre-built apis address common perceptual tasks such as vision video and natural language they are ready to use without any ml expertise or model development effort if your developers and data scientists want to build custom models with your own training data while spending minimal time coding then automl is your choice automl provides a codeless solution to enable you to focus on business problems instead of the underlying model architecture and ml provisioning if your ml engineers and data scientists want full control of ml workflow vertex ai custom training lets you train and serve custom models with code on vertex workbench



we've already explored bigquery ml so in the videos that follow we'll explore the other three options in more detail.

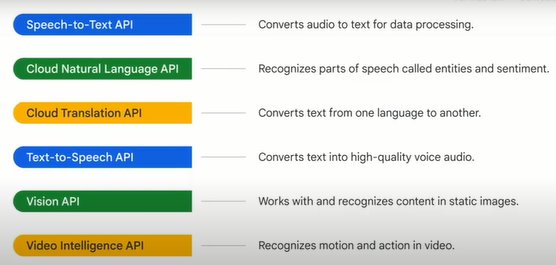
**Pre-builts APIs.**

good machine learning models require lots of high quality training data you should aim for hundreds of thousands of records to train a custom model if you don't have that kind of data pre-built apis are a great place to start pre-built apis are offered as services in many cases they can act as building blocks to create the application you want without the expense or complexity of creating your own models they save the time and effort of building curating and training a new data set so you can quickly move to predictions so what are some of the pre-built apis

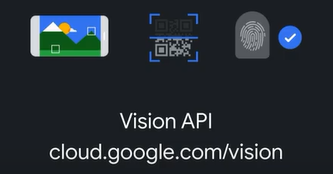


let's explore a short list:

the speech-to-text api converts audio to text for data processing the cloud natural language api recognizes parts of speech called entities and sentiment the cloud translation api converts text from one language to another the text-to-speech api converts text into high quality voice audio the vision api works with and recognizes content in static images and the video intelligence api recognizes motion and action in video google has already done a lot of work to train these models using google data sets



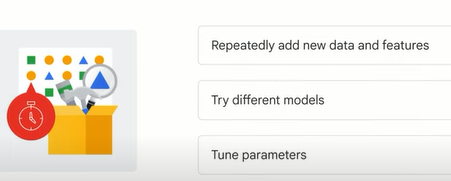
for example the vision api is based on google's image data sets the speech-to-text api is trained on youtube captions and the translation api is built on google translations you'll recall that how well a model is trained depends on how much data is available to train it as you might expect google has a lot of images text and ml researchers to train its pre-built models this means less work for you let's try out the vision api in a browser start by navigating to cloud.google.com/vision in chrome and then scroll down to try the api by uploading an image you can actually experiment with each of the ml apis in a browser



when you're ready to build a production model you'll need to pass a json object request to the api and parse what it returns.

**AutoML.**

to understand automl which is short for automated machine learning let's briefly look at how it was built if you've worked with ml models before you know that training and deploying ml models can be extremely time consuming because you need to repeatedly add new data and features try different models and tune parameters to achieve the best result



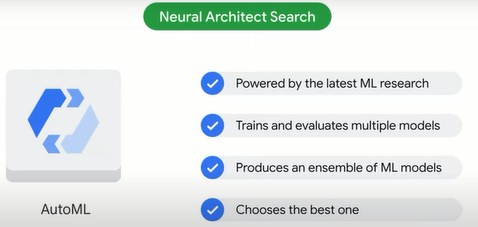
to solve this problem when automl was first announced in january 2018 the goal was to automate machine learning so data scientists didn't have to start the process from scratch (desde cero).

but how could this be done well machine learning is similar to human learning it all starts with gathering the right information for automl two technologies are vital:

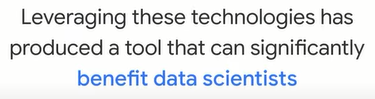
**the first is transfer learning** with transfer learning you build a knowledge base in the field you can think of this like gathering lots of books to create a library.

transfer learning is a powerful technique that lets people with smaller data sets or less computational power achieve state-of-the-art results by taking advantage of pre-trained models that have been trained on similar larger data sets because the model learns via transfer learning it doesn't have to learn from scratch so it can generally reach higher accuracy with much less data and computation time than models that don't use transfer learning.

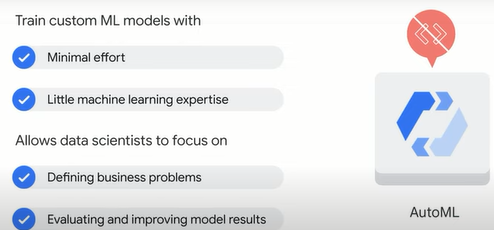
**the second technology is neural**. architect search the goal of neural architect search is to find the optimal model for the relevant project think of this like finding the best book in the library to help you learn what you need to automl is powered by the latest machine learning research so although a model performs training the automl platform actually trains and evaluates multiple models and compares them to each other this neural architecture search produces an ensemble of ml models and chooses the best one



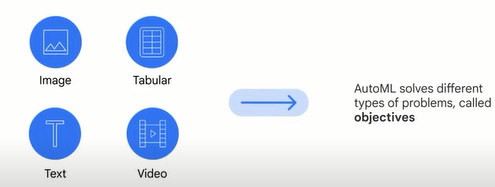
leveraging these technologies has produced a tool that can significantly benefit data scientists



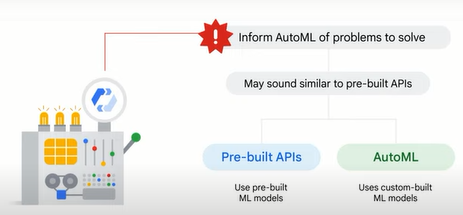
one of the biggest benefits is that it's a no code solution that means it can train high quality custom machine learning models with minimal effort and requires little machine learning expertise this allows data scientists to focus their time on tasks like defining business problems or evaluating and improving model results



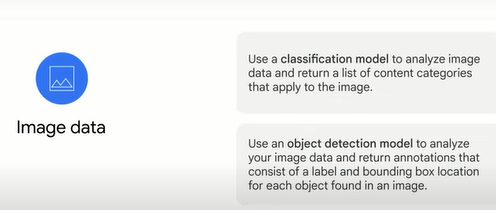
others might find automl useful as a tool to quickly prototype models and explore new datasets before investing in development this might mean using it to identify the best features in a dataset for example so how does it work automl supports four types of data image tabular text and video for each data type automl solves different types of problems called objectives



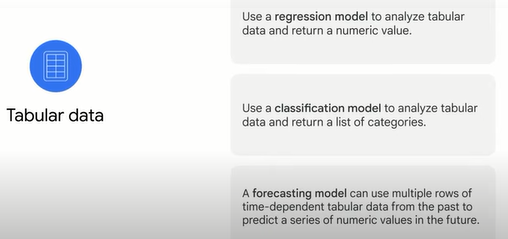
to get started upload your data into automl it can come from cloud storage bigquery or even your local machine from there inform automl of the problems you want to solve some problems may sound similar to those mentioned in pre-built apis however the major difference is that pre-built apis use pre-built machine learning models but automl uses custom built models in automl you use your own data to train the machine learning model and then apply the trained model to predict your goal.



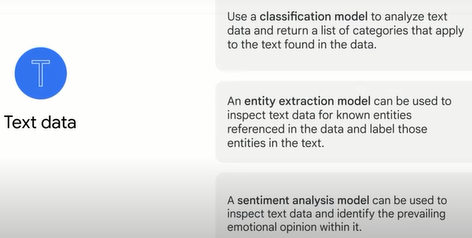
**for image data** you can use a classification model to analyze image data and return a list of content categories that apply to the image for example you could train a model that classifies images as containing a dog or not containing a dog or you could train a model to classify images of dogs by breed can also use an object detection model to analyze your image data and return annotations that consist of a label and bounding box location for each object found in an image for example you could train a model to find the location of the dogs in image data



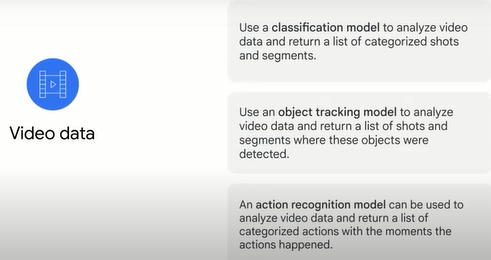
**for tabular data** you can use a regression model to analyze tabular data and return a numeric value for example you could train a model to estimate a house's value or rental price based on a set of factors such as location size of the house and number of bedrooms you can use a classification model to analyze tabular data and return a list of categories for example you could train a model to classify different types of land into high median and low potentials for commercial real estate and a forecasting model can use multiple rows of time-dependent tabular data from the past to predict a series of numeric values in the future for example you could use the historical plus the economic data to predict what the housing market will look like in the next five years



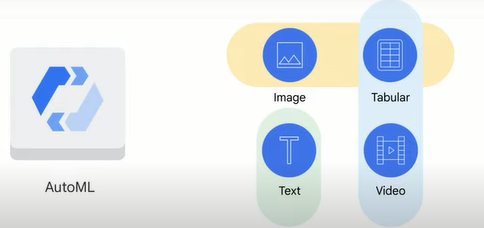
**for text data** you can use a classification model to analyze text data and return a list of categories that's applied to the text found in the data for example you can classify customer questions and comments to different categories and then redirect them to corresponding departments an entity extraction model can be used to inspect text data for known entities referenced in the data and label those entities in the text for example you can label a social media post in terms of predefined entities such as time location and topic this can help with online search similar to the concept of a hashtag but created by a machine and a sentiment analysis model can be used to inspect text data and identify the prevailing emotional opinion within it especially to determine a writer's comment as positive negative or neutral



and finally **for video data** you can use a classification model to analyze video data and return a list of categorized shots and segments for example you could train a model that analyzes video data to identify whether the video is of a soccer baseball basketball or football game you can use an object tracking model to analyze video data and return a list of shots and segments where these objects were detected for example you could train a model that analyzes video data from soccer games to identify and track the ball and an action recognition model can be used to analyze video data and return a list of categorized actions with the moments the actions happened for example you could train a model that analyzes video data to identify the action moments involving a soccer goal a golf swing a touchdown or a high five

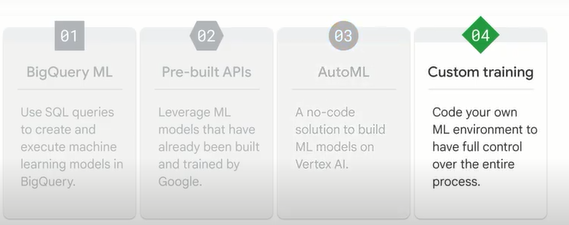


in reality you may not be restricted to just one data type and one objective but instead need to combine multiple data types and different objectives to solve a business problem automl is a powerful tool that can help across these different data types and objectives.



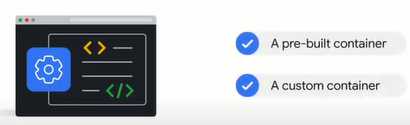
**Custom training.**

we've explored the options google cloud provides to build machine learning models using bigquery ml pre-built apis and automl now let's take a look at the last option custom training



if you want to code your machine learning model you can use this option by building a custom training solution with vertex ai workbench.

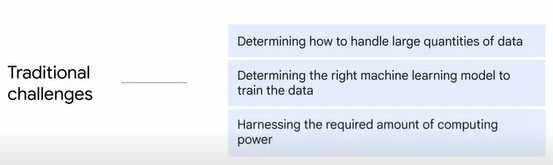
**workbench** is a single development environment for the entire data science workflow from exploring to training and then deploying a machine learning model with code before any coding begins you need to determine what environment you want your ml training code to use there are two options **a pre-built container or a custom container**



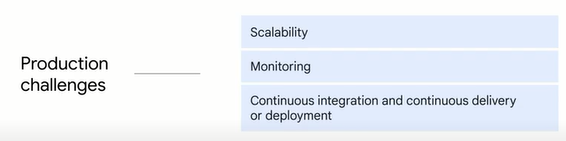
imagine that a container is a kitchen a pre-built container would represent a fully furnished room with cabinets and appliances which represent the dependencies that includes all the cookware which represents the libraries you need to make a meal so if your ml training needs a platform like tensorflow pytorch scikit-learn or xgboost and python code to work with the platform a pre-built container is probably your best solution a custom container alternatively is like an empty room with no cabinets appliances or cookware you define the exact tools you need to complete the job.

**Vertex AI.**

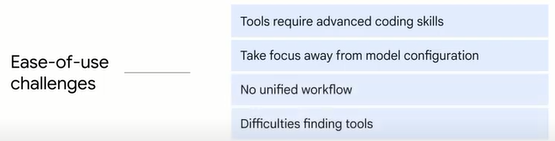
for years google has invested time and resources into developing big data and ai google has developed key technologies and products from its roots back in the development of psychic learn back in 2007 to vertex ai today as an ai first company google has applied ai technologies to many of its products and services like gmail google maps google photos and google translate just to name a few but developing these technologies doesn't come without challenges especially when it involves developing machine learning models and putting them into production some traditional challenges include determining how to handle large quantities of data determining the right machine learning model to train the data and harnessing the required amount of computing power



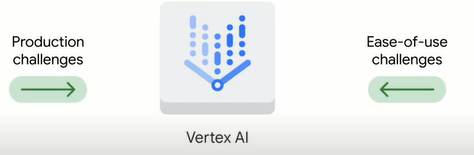
then there are challenges around getting ml models into production production requires scalability monitoring and continuous integration and continuous delivery or deployment these challenges can make projects fail



in fact according to gartner only half of enterprise ml projects get past the pilot phase and finally there are ease of use challenges many tools on the market require advanced coding skills which can take a data scientist's focus away from model configuration and without a unified workflow data scientists often have difficulties finding tools



google's solution to many of the production and ease of use challenges is vertex ai a unified platform that brings all the components of the machine learning ecosystem and workflow together



so what exactly does a unified platform mean?

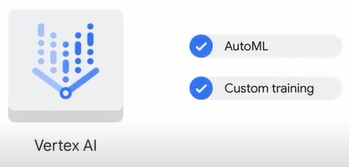
in the case of vertex ai it means having one digital experience to create deploy and manage models over time and at scale.



For example during the data readiness stage users can upload data from wherever it's stored cloud storage bigquery or a local machine then during the feature readiness stage users can create features which are the process data that will be put into the model and then share them with others using the feature store after that it's time for training and hyper parameter tuning this means that when the data is ready users can experiment with different models and adjust hyper parameters and finally during deployment and model monitoring users can set up the pipeline to transform the model into production by automatically monitoring and performing continuous improvements.



And to refer back to the different options we explored earlier vertex ai allows users to build machine learning models with either automl a codeless solution or custom training a code-based solution automl is easy to use and lets data scientists spend more time turning business problems into ml solutions while custom training enables data scientists to have full control over the development environment and process

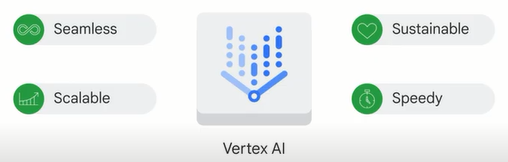


being able to perform such a wide range of tasks in one unified platform has many benefits this can be summarized with 4s's it's **seamless** vertex i provides a smooth (fluida,tranquila) user experience from uploading preparing data all the way to model training and production

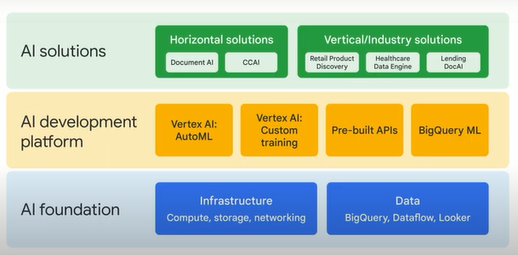
it's **scalable** the machine learning operations ml ops provided by vertex ai helps to monitor and manage the ml production and therefore scale the storage and computing power automatically

it's **sustainable** all of the artifacts and features created using vertex ai can be reused and shared

and it's **speedy** vertex i produces models that have eighty percent fewer lines of code than competitors.



**AI Solutions.**

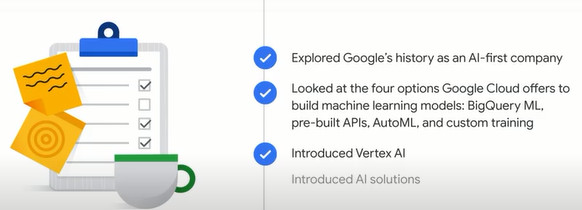


now that you've explored the four different options available to create machine learning models with google cloud let's take a few minutes to look at google cloud's artificial intelligence solution portfolio it can be visualized with three layers the bottom layer is the ai foundation and includes the google cloud infrastructure and data the middle layer represents the ai development platform which includes the four ml options you just learned about automl and custom training which are offered through vertex ai pre-built apis and bigquery ml the top layer represents ai solutions for which there are two groups horizontal solutions and industry solutions horizontal solutions usually apply to any industry that would like to solve the same problem examples include document ai and ccai document ai uses computer vision and optical character recognition along with natural language processing to create pre-trained models to extract information from documents the goal is to increase the speed and accuracy of document processing to help organizations make better decisions faster while reducing costs another example of a horizontal solution is contact center ai or ccai the goal of ccai is to improve customer service in contact centers through the use of artificial intelligence it can help automate simple interactions assist human agents unlock caller insights and provide information to answer customer questions the second group is vertical or industry solutions this represents solutions that are relevant to specific industries examples include retail product discovery which gives retailers the ability to provide google quality search and recommendations on their own digital properties helping to increase conversions and reduce search abandonment google cloud healthcare data engine which generates healthcare insights and analytics with one end-to-end solution and lending.ai which aims to transform the home loan experience for borrowers and lenders by automating mortgage document processing you can learn more about google cloud's growing list of ai solutions at cloud.google.com/solutions/ai.

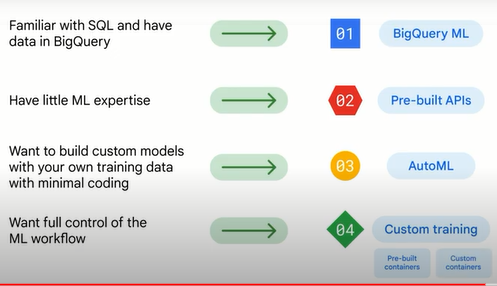


**Summary.**

we've covered a lot of information in this section of the course let's do a quick recap to start we explored google's history as an ai first company from there we looked at the four options google cloud offers to build machine learning models this includes bigquery ml pre-built apis automl and custom training next we introduced vertex ai a tool that combines the functionality of automl which is codeless and custom training which is code-based to solve production and ease-of-use problems

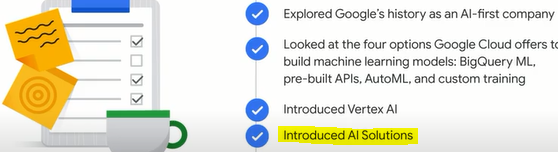


you'll recall that selecting the best option will depend on your business needs and ml expertise if your data engineers data scientists and data analysts are familiar with sql and already have your data in bigquery bigquery ml lets you develop sql based models if your business users or developers have little ml experience using pre-built apis is likely the best choice pre-built apis address common perceptual tasks such as vision video and natural language they are ready to use without any ml expertise or model development effort if your developers and data scientists want to build custom models with your own training data while spending minimal time coding then automl is your choice automl provides a codeless solution to enable you to focus on business problems instead of the underlying model architecture and ml provisioning if your ml engineers and data scientists want full control of the ml workflow vertex ai custom training lets you train and serve custom models with code on vertex workbench using pre-built containers you can leverage popular ml libraries such as tensorflow and pytorch alternatively you can build a custom container from scratch

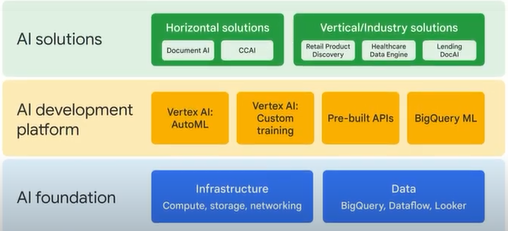


and finally we introduced google cloud ai solutions

.



the solutions are built on top of the four ml development options to meet both horizontal and vertical market needs



**Quiz**

1.You work for a video production company and want to use machine learning to categorize event footage, but want to train your own ML model. Which option can help you get started?

BigQuery ML

AutoML

**Pre-built APIs**

Custom training

2.Your company has a lot of data, and you want to train your own machine model to see what insights ML can provide. Due to resource constraints, you require a codeless solution. Which option is best?

BigQuery ML

**AutoML**

Pre-built APIs

Custom training

3. You work for a global hotel chain that has recently loaded some guest data into BigQuery. You have experience writing SQL and want to leverage machine learning to help predict guest trends for the next few months. Which option is best?

**BigQuery ML**

AutoML

Pre-built APIs

Custom training

4. Which code-based solution offered with Vertex AI gives data scientists full control over the development environment and process?

AutoML

AI Platform

**Custom training**

AI Solutions

5.Which Google Cloud product lets users create, deploy, and manage machine learning models in one unified platform?

Document AI

AI Platform

TensorFlow

**Vertex AI**