Google Developer Groups

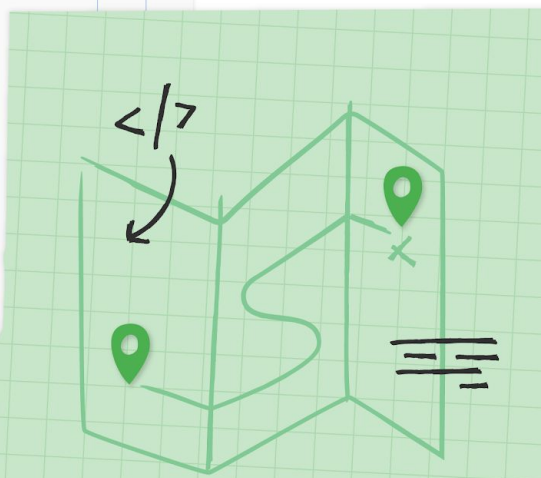
Road to Google Developers Certification

Google Cloud

Week 1

Professional Machine Learning Engineer

GDG Lawrence



```
udies,
tudies.
erByOrg
filterBy
hStatus
```

```
let filteredStudies = studies.filter(study => {
  const matchOrg = filterByOrg ? study.lead_organization
  const matchStatus = filterByStatus ? study.status === fi
  if (matchOrg && matchStatus) {
    return true
  }
  return false
```

Google Developer Groups

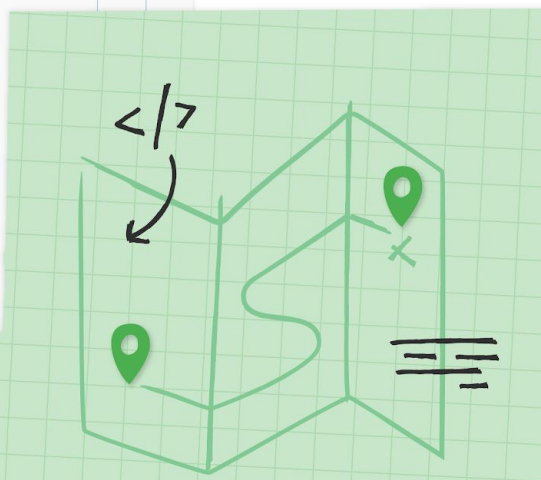
Road to Google Developers Certification

Google Cloud

Thanks to all co-hosting chapters

GDG Cloud Boston
GDG Capital Region
GDG San Antonio
GDG Portland

GDG Ocala
GDG Tampa Bay
GDG Space Coast
GDG Modesto
GDG Toledo



```
udies,
tudies.
erByOrg
filterBy
hStatus
```

```
let filteredStudies = studies.filter(study => {
  const matchOrg = filterByOrg ? study.lead_organization
  const matchStatus = filterByStatus ? study.status === fi
  if (matchOrg && matchStatus) {
    return true
  }
  return false
```

Professional ML Engineer

Learning Journey

Google Developer Groups
Road to Google Developers Certification
Google Cloud

Pre-work	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Exam weeks
Exam guide review sessions ➪	kick-off session	exam guide review session	exam guide review session	exam guide review session	exam guide review session	exam guide review session + wrap up	**Self study (and potential exam)**
Review the Professional ML Engineer Exam Guide **Review the** Professional ML Engineer Sample Questions **Go through:** Google Cloud Platform Big Data and Machine Learning Fundamentals	**Complete course:** How Google does Machine Learning Launching into Machine Learning	**Complete course:** TensorFlow on Google Cloud Feature Engineering	**Complete course:** Machine Learning in the Enterprise	**Hands On Lab Practice:** Production Machine Learning Systems Computer Vision Fundamentals with Google Cloud	**Complete course:** Natural Language Processing on Google Cloud Recommendation Systems on GCP	**Complete course:** ML Ops - Getting Started ML Pipelines on Google Cloud **Check Readiness:** Professional ML Engineer Sample Questions **Hands On Lab Practice:** Perform Foundational Data, ML, and AI Tasks in Google Cloud (Skill Badge) - 7hrs Build and Deploy ML Solutions on Vertex AI (Skill Badge) - 8hrs	

Swag Raffle!!!

Stick ‘til the end!

Google Cloud Unisex Onyx Tee

Google Cloud Striped Sock

Google Cloud Cap

Google Cloud Clay Mug

Google Cloud

How Google Does Machine Learning

Machine learning is a way to use standard algorithms to derive predictive insights from data and make repeated decisions.

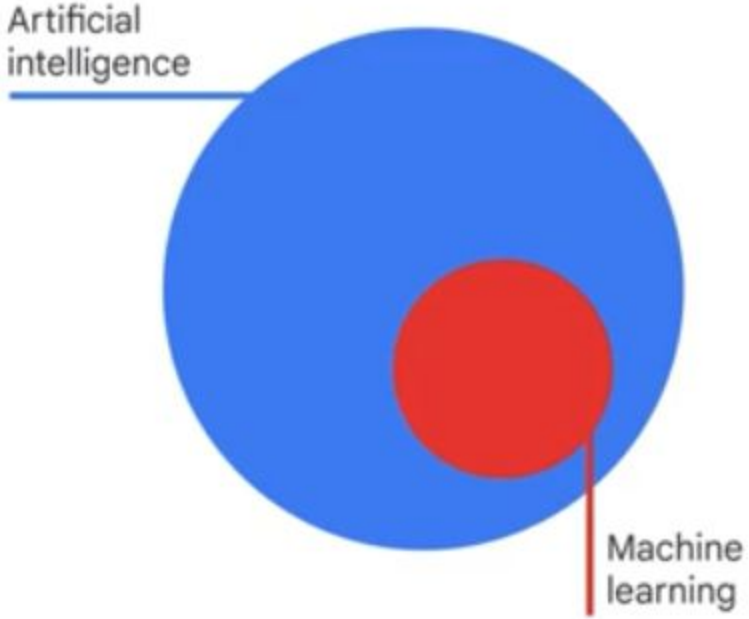
Data

Algorithm

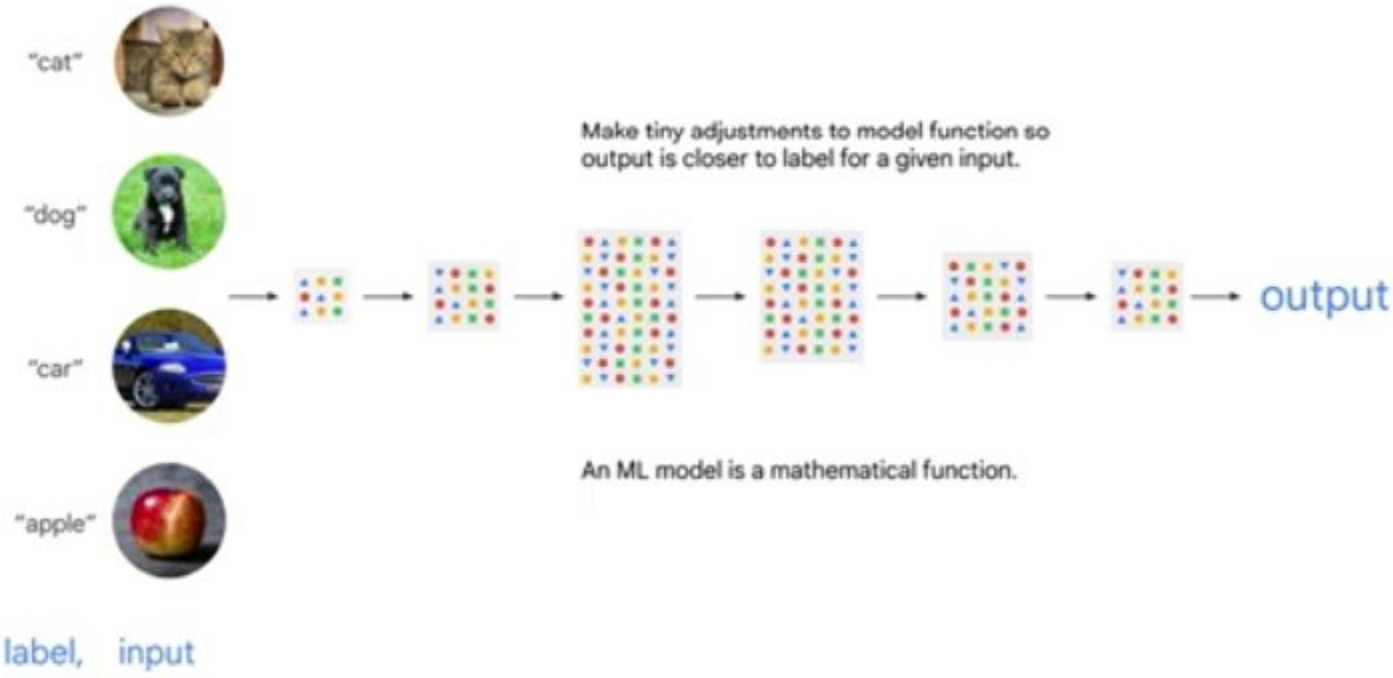
Predictive insight

Decision

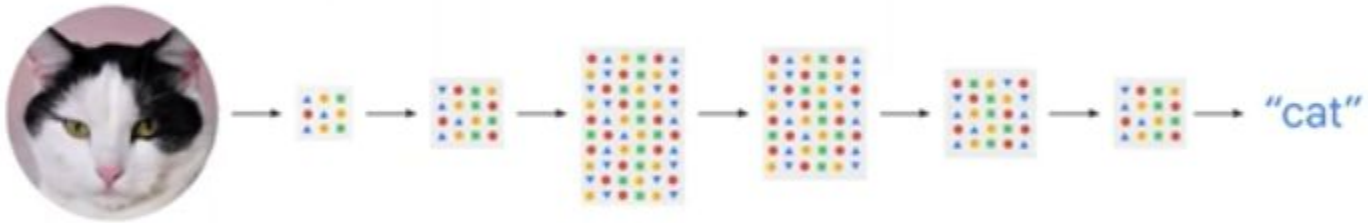
Artificial Intelligence is a discipline; machine learning is a specific way of solving AI problems.

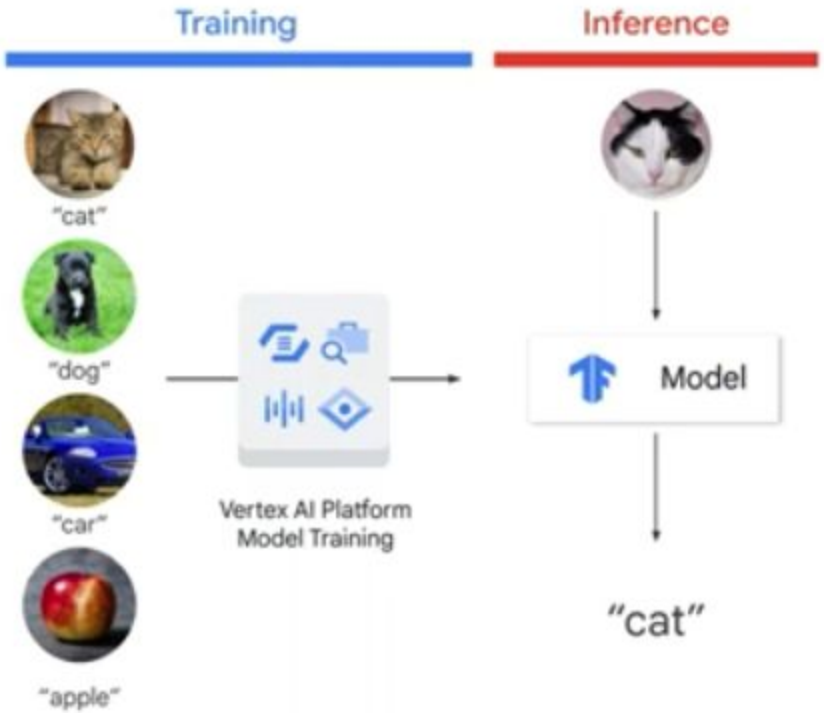


Stage 1: Train an ML model with examples



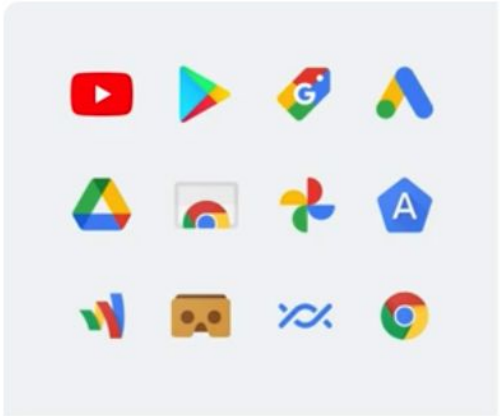
Stage 2: Predict with a trained model



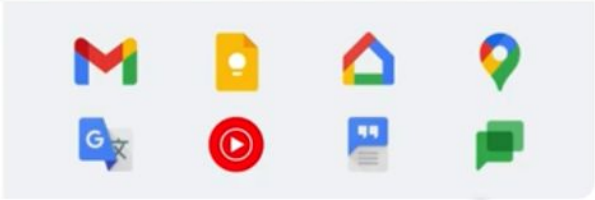


Focus on both the training and inference stages of ML

Google has more than
10,000 deep learning models



Google infuses
Machine Learning into
almost all its products.



Deep learning has come a long way in just the past few years

Google Photos

illustrates how far ML has come.

Google Translate

is a combination of several models.

Gmail

Smart Reply Inbox
20% of all responses sent on mobile.

Pre-trained models

There are pre-trained machine learning services available on Google Cloud

Custom ML models

Vertex AI

TensorFlow

Pre-trained ML Models

Vision API

Speech API

Jobs API

Translation API

Natural Language API

Video Intelligence API

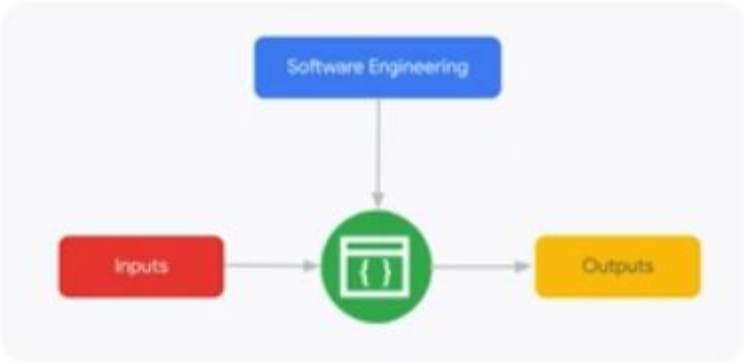
1. Collecting data is often the longest and hardest part of an ML project, and the one most likely to fail

2. Manual analysis helps you fail fast and try new ideas in a more agile way

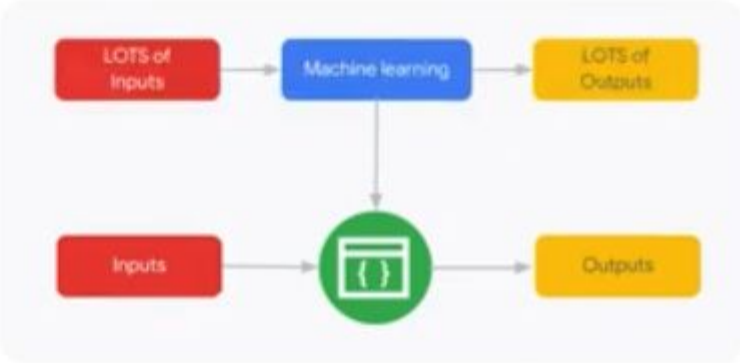
3. To build a good ML model, you have to know your data

4. ML is a journey towards automation and scale

Software engineers write program rules



Machine learning figures out program rules



Avoid these top 10 ML pitfalls

Defining KPIs · Collecting data · Integration · Infrastructure · Optimizing ML

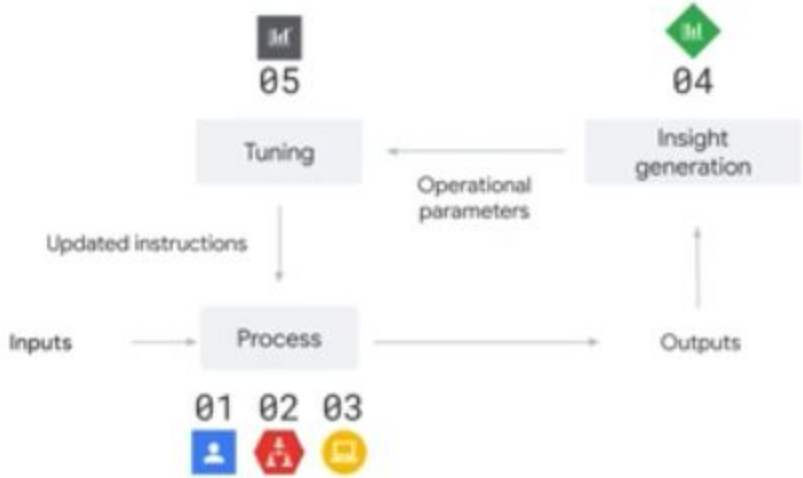
01 ML requires just as much software infrastructure
02 No data collected yet
03 Assume the data is ready for use
04 Keep humans in the loop
05 Product launch focused on the ML algorithm
06 ML optimizing for the wrong thing
07 Is your ML improving things in the real world
08 Using a pre-trained ML algorithm vs building your own
09 ML algorithms are trained more than once
10 Trying to design your own perception or NLP algorithm

Path to ML: The 5 phases

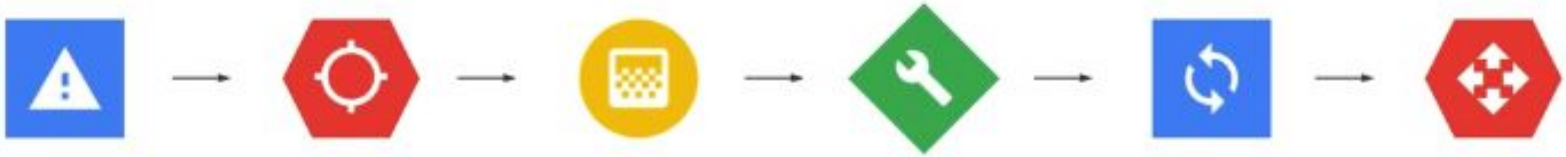
How change happens in phases:

- 01 Individual contributor
- 02 Delegation
- 03 Digitization
- 04 Big data and analytics
- 05 Machine learning

General feedback loop

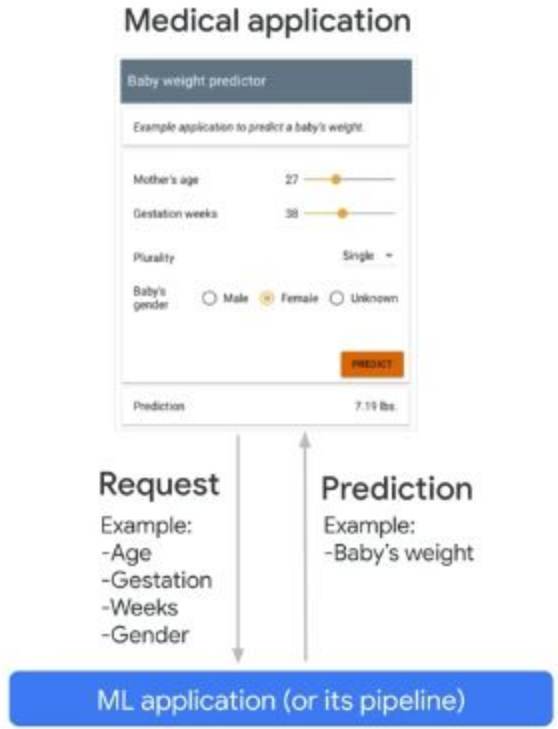


To build a machine learning model for production



1	2	3	4	5	6
Identify goal	Acquire + explore data + prepare data	Build a model	Train the model	Evaluate the model	Deploy + monitor + maintain

ML application generates a REST service for use by a medical application



What is there to unify?

Dataset is

- Created
- Ingested
- Analyzed
- Cleaned (ETL or ELT)

Model is

- **Trained,** which includes experimentation and hypothesis. testing, and hyperparameter tuning.
- **Versioned** and rebuilt when there is new data, on a schedule, or when the code changes (ML Ops).
- **Evaluated** and compared to existing model versions.
- **Deployed** and used for online and batch predictions.

Source: Giving Vertex AI, the New Unified ML Platform on Google Cloud, a Spin, by Lak Lakshmanan

Choose a training method

AutoML

- Create and train a model with minimal technical effort.
- Quickly prototype models or explore datasets before developing in a custom training application.

Custom training

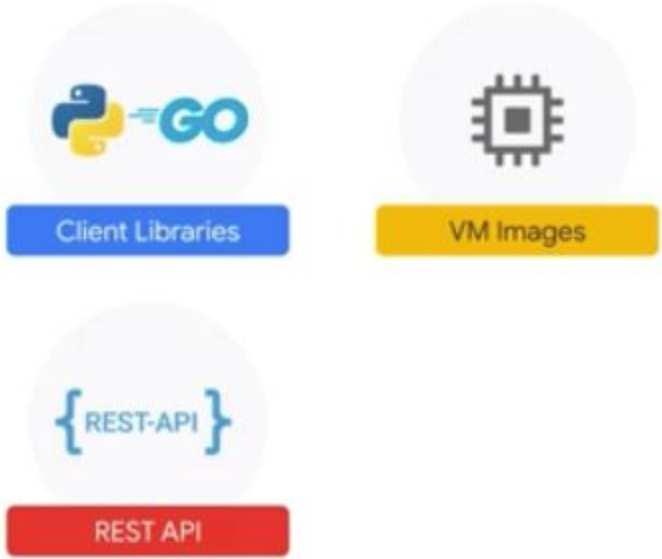
- Create a training application optimized for your targeted outcome.
- Maintain complete control over training application functionality.
  - Target any objective, use any algorithms, develop your own loss functions or metrics, or make other customizations.


Vertex AI

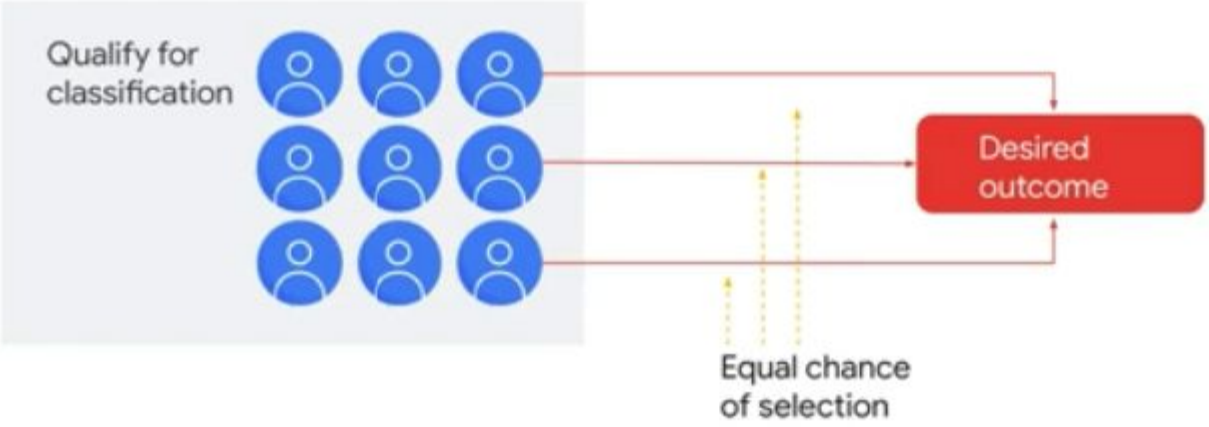
Vertex AI provides client libraries for some languages to help you make calls to the Vertex AI API.

Alternatively, you can use the Google API Client Libraries to access the Vertex AI API by using other languages.

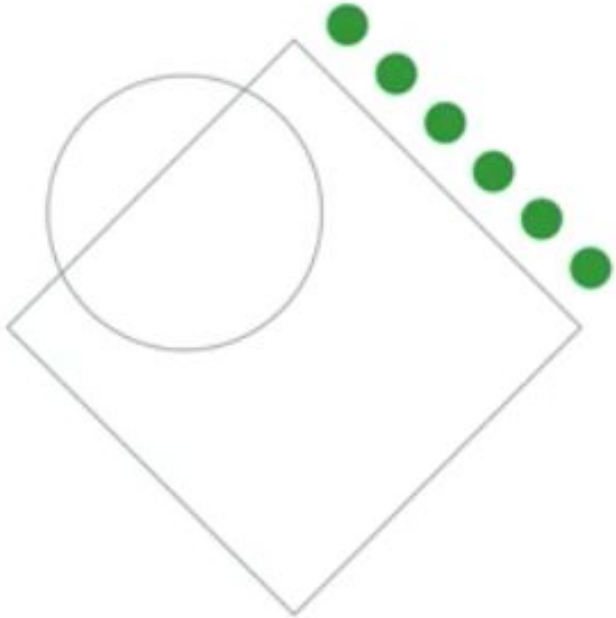
Tools to interact with Vertex AI



The equality of opportunity approach strives to give individuals an equal chance of desired outcome



Introduction to automated machine learning (AutoML) using Vertex AI



Machine learning versus statistics

Machine learning

Lots of data. Keep outliers and build models for them.

Statistics

"I've got all the data I'll ever get." Throw away outliers.

Factors

- Data requirement
- Accuracy
- Training time
- Hardware dependency

Deep learning

- Requires large data
- Provides high accuracy
- Takes longer to train
- Requires GPU to train properly

Machine learning

- Can train on lesser data
- Gives lesser accuracy
- Takes less time to train
- Trains on CPU

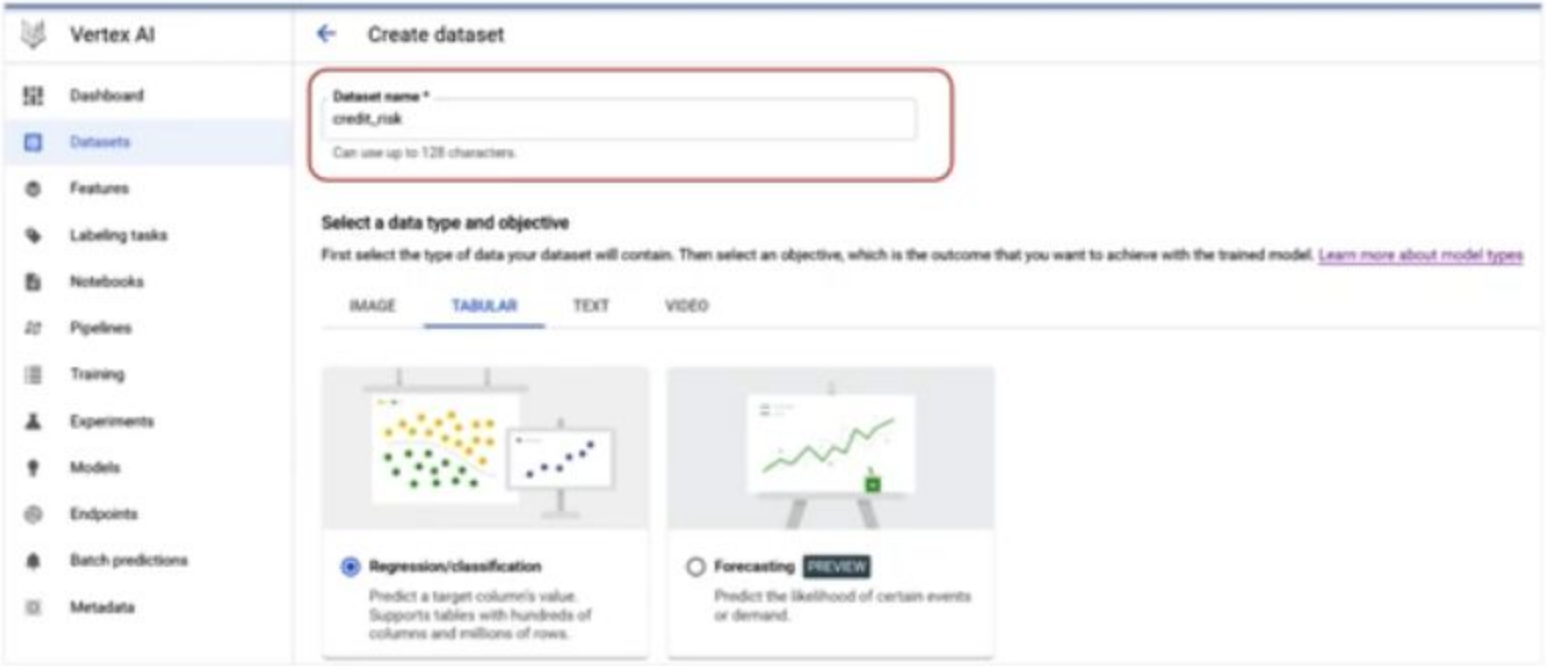
Dataset

```
[1] import pandas as pd
```

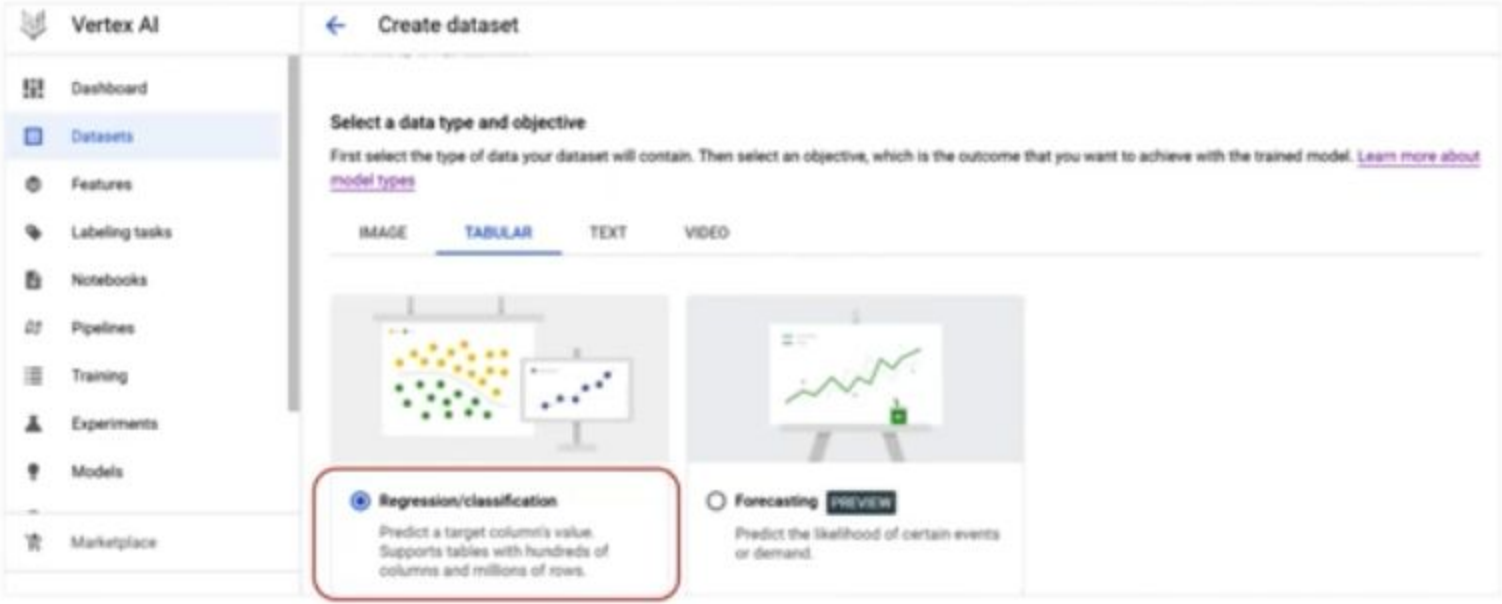
```
df = pd.read_csv('consumer_spend.csv')
data.head()
```

	Graduated	Profession	Work_Experience	Family_Size	Spending_Score
0	No	Healthcare	1.0	4.0	Low
1	Yes	Engineer	NaN	3.0	Average
2	Yes	Engineer	1.0	1.0	Low
3	Yes	Lawyer	0.0	2.0	High
4	Yes	Entertainment	NaN	6.0	High

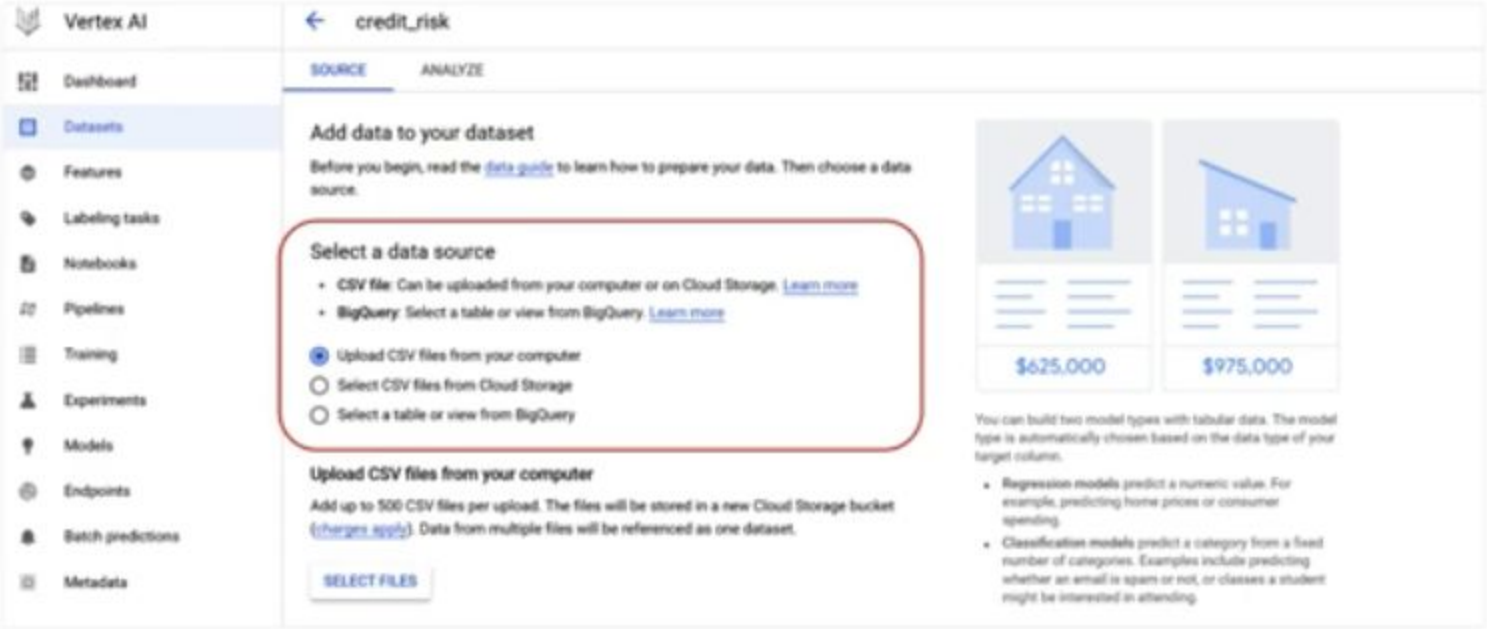
Step 1: Create a dataset



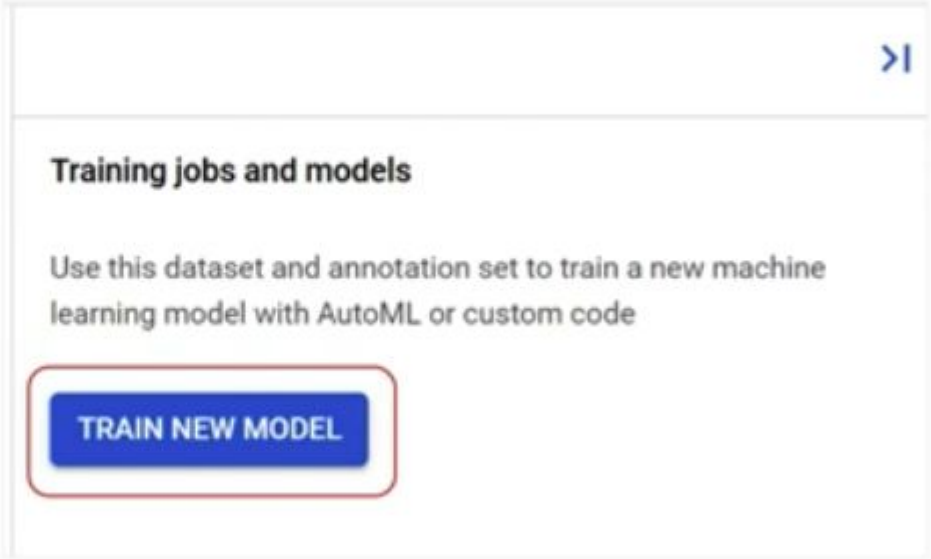
Step 2: Select a datatype and objective



Step 3: Upload the data



Step 4: Train a new model



Vertex AI finished training model "credit_risk_202182831655" Inbox ×

Vertex AI
to me

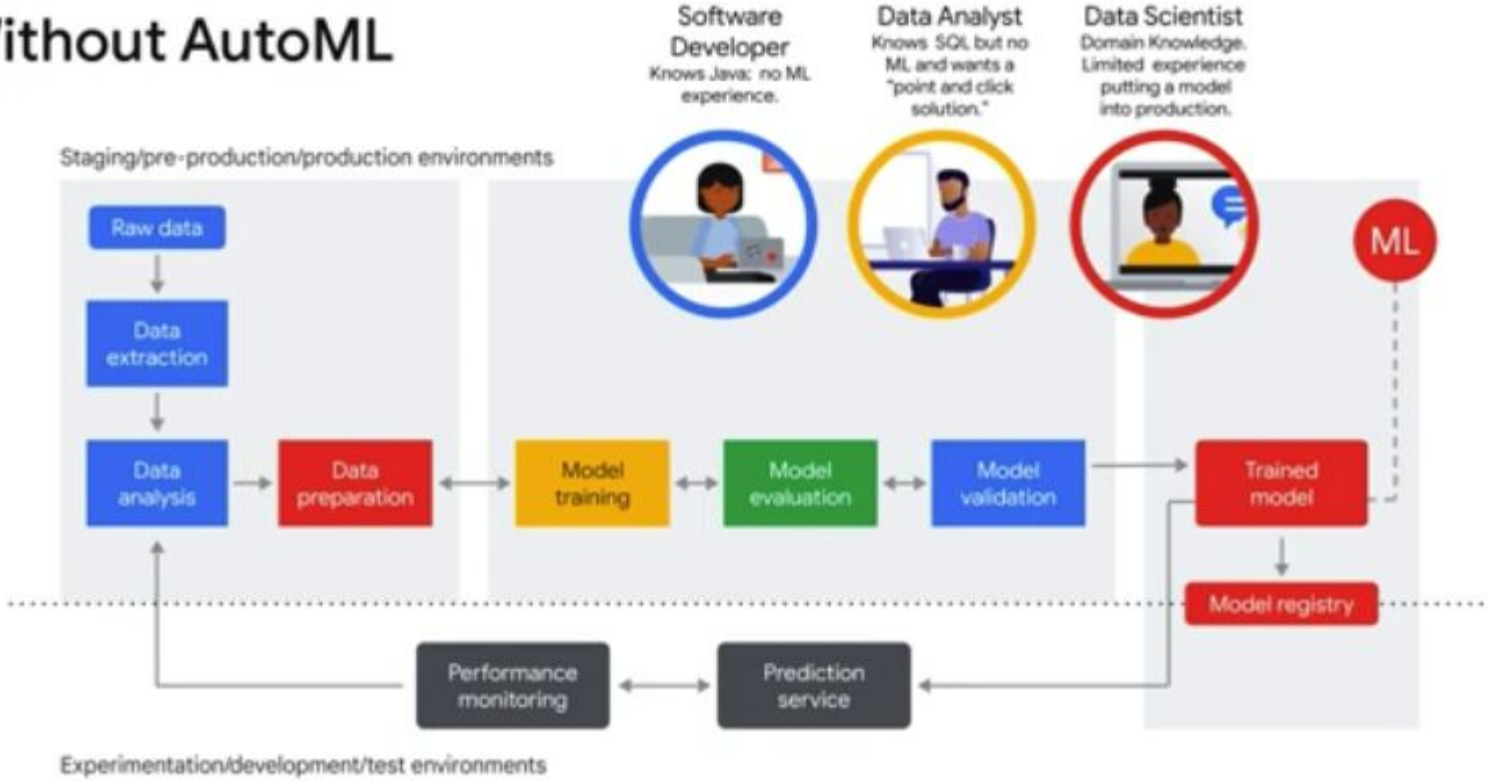
Hello Vertex AI Customer,

Vertex AI finished training model "credit_risk_202182831655".
Additional Details:
Operation State: Succeeded
Resource Name:
projects/663413318684/locations/us-central1/trainingPipelines/4923316201041428480

To continue your progress, go back to your training pipeline using
https://pantheon.corp.google.com/vertex-ai/models?project=cloud-training-demos

Sincerely,
The Google Cloud AI Team

Without AutoML



Where Vertex AI fits in the ML workflow

You can use Vertex AI to manage the following stages in the ML workflow:

Create a dataset and upload data.

Train an ML model on your data:

- Train the model
- Evaluate model accuracy
- Tune hyperparameters (custom training only)

Upload and store your model in Vertex AI.

Deploy your trained model to an endpoint for serving predictions.

Send prediction requests to your endpoint.

Specify a prediction traffic split in your endpoint.

Manage your models and endpoints.

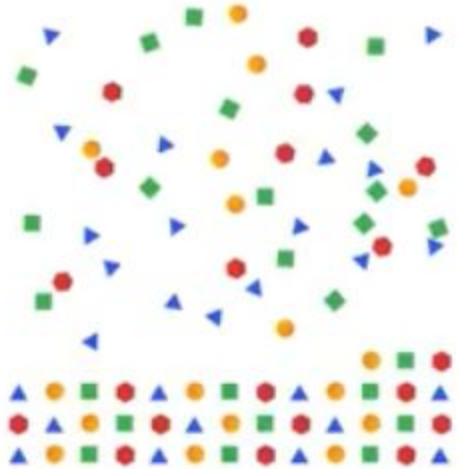
When to use AutoML and when to use custom training

	AutoML	Custom training
Data science expertise needed	No.	Yes, to develop the training application and also to do some of the data preparation like feature engineering.
Programming ability needed	No, AutoML is codeless.	Yes, to develop the training application
Time to trained model	Lower. Less data preparation is required, and no development is needed.	Higher. More data preparation is required, and training application development is needed.
Limits on machine learning objectives	Yes. You must target one of AutoML's predefined objectives.	No.
Can manually optimize model performance with hyperparameter tuning	No. AutoML does some automated hyperparameter tuning, but you can't modify the values used.	Yes. You can tune the model during each training run for experimentation and comparison.

How AutoML Tables uses your dataset

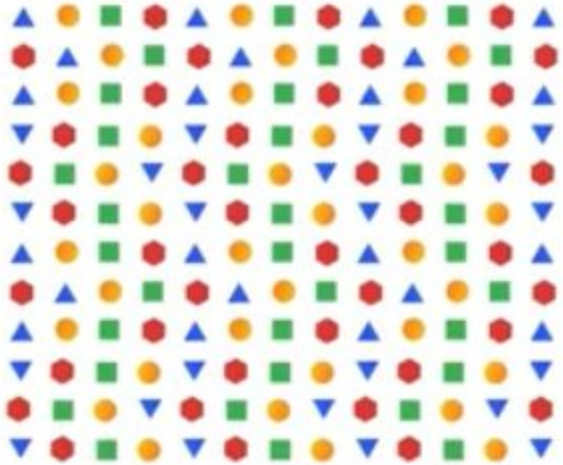
Training set

- The vast majority of your data should be here.



Validation set

- The validation set is sometimes referred to as the "dev" set, and is also used in the training process.



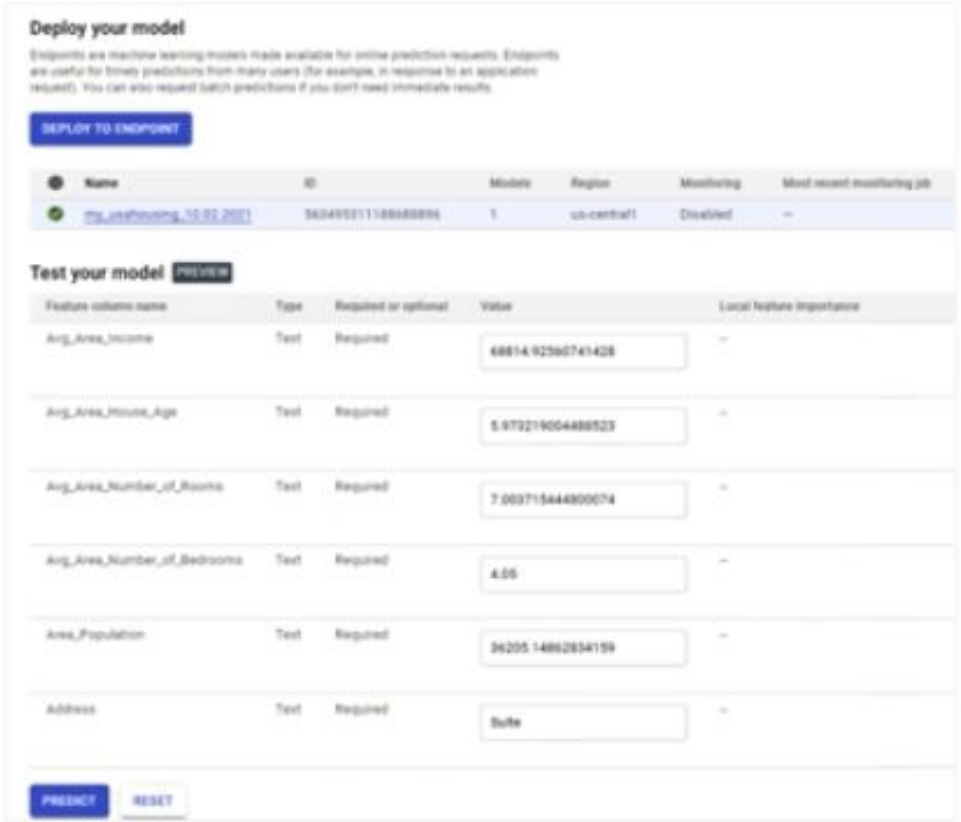
Test set

- The test set is used after the training process is complete.

Test your model

- After evaluating your model metrics, you can test your model with new data.
- See if the model's predictions match your expectations.
- If not, you may need to continue improving your model's performance.

Deploy your model



Deploy your model and make online predictions

Batch prediction

- Allows you to make many prediction requests at once.
- Is asynchronous (the model won't return a CSV file or BigQUery Table until it processes all prediction requests).

Online prediction

- Deploy your model to make it available for prediction requests using a REST API.
- Is synchronous (the model will quickly return a prediction, but only accepts one prediction request per API call).

Working with BigQuery ML

01

Dataset

02

Create/ train

```
CREATE MODEL `BigQuery
ML_tutorial.sample_model`
OPTIONS(model_type='logistic_reg') AS
SELECT
```

03

Evaluate

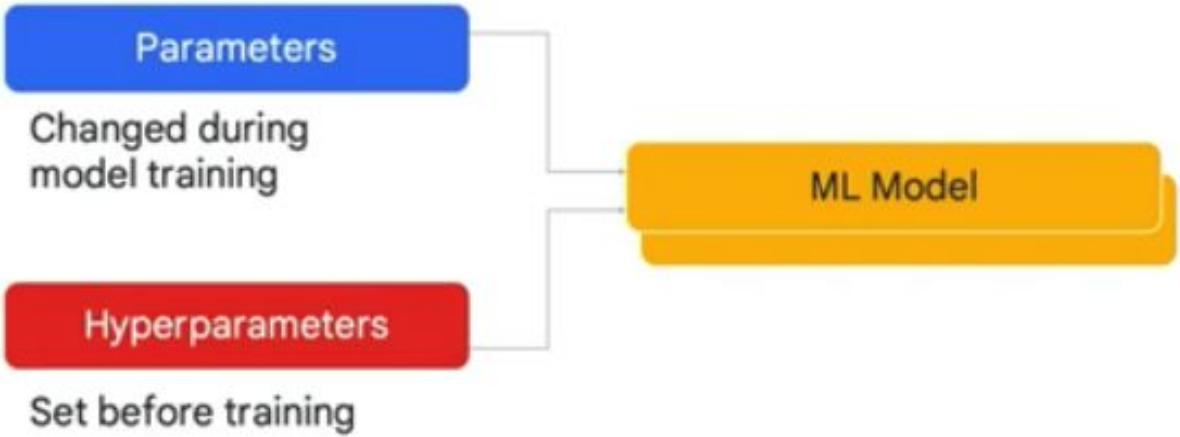
```
FROM
  ML.EVALUATE(MODEL `BigQuery
ML_tutorial.sample_model`,
    TABLE eval_table)
```

04

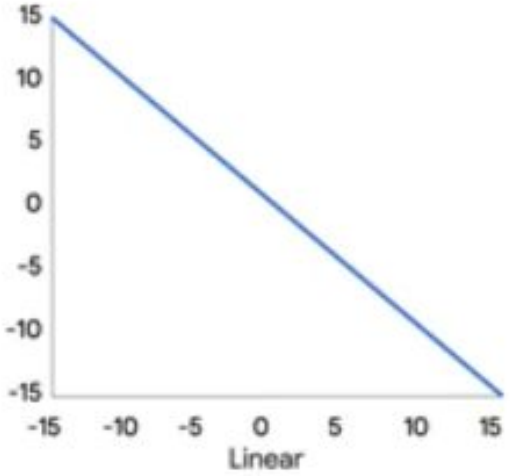
Predict/ classify

```
FROM
    ML.PREDICT(MODEL `BigQuery
ML_tutorial.sample_model`,
      table game_to_predict) ) AS
predict
```

ML models are mathematical functions with parameters and hyperparameters



Linear models have two types of parameters: Bias and weight

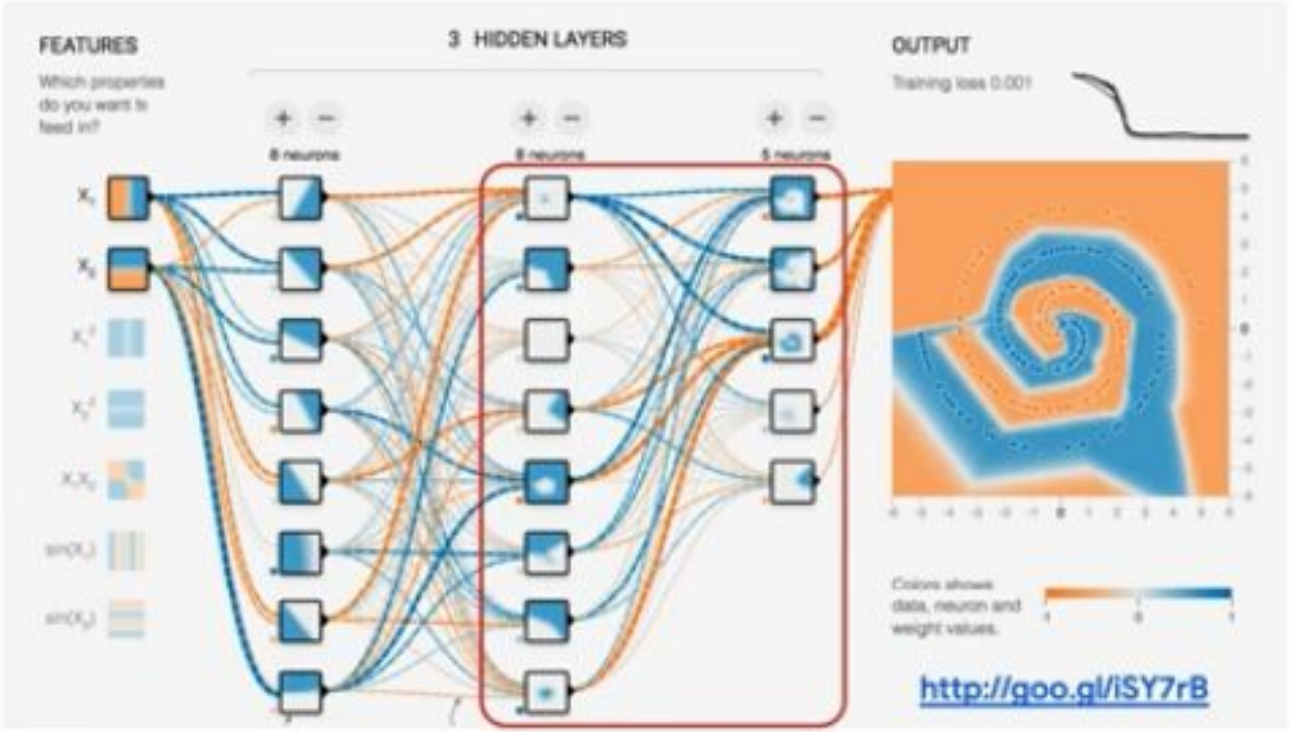


Output Bias Term Input Weight

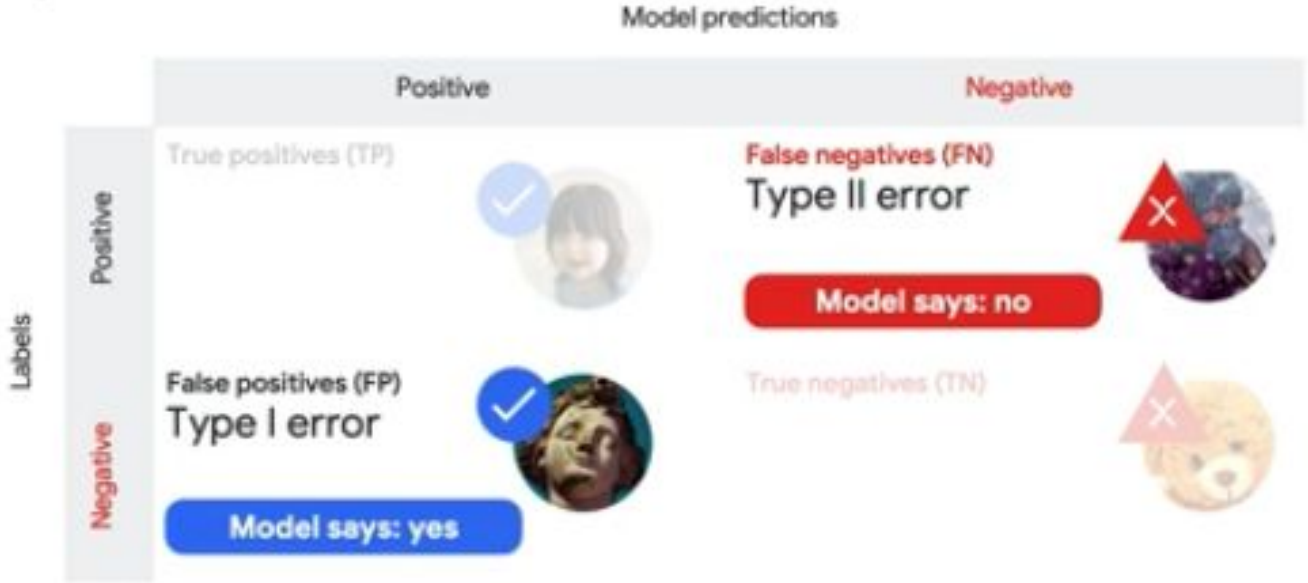
$$y = b + x \times m$$

Model parameters

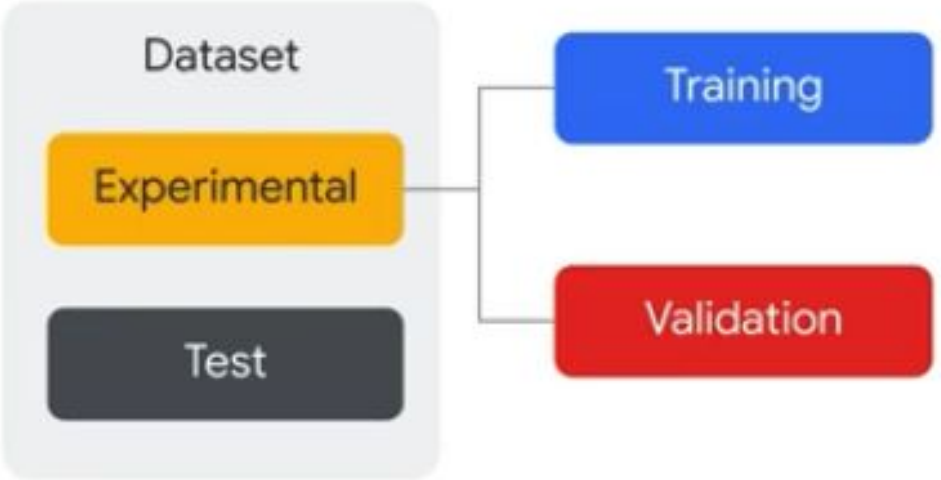
More hidden layers leads to more hierarchies of features



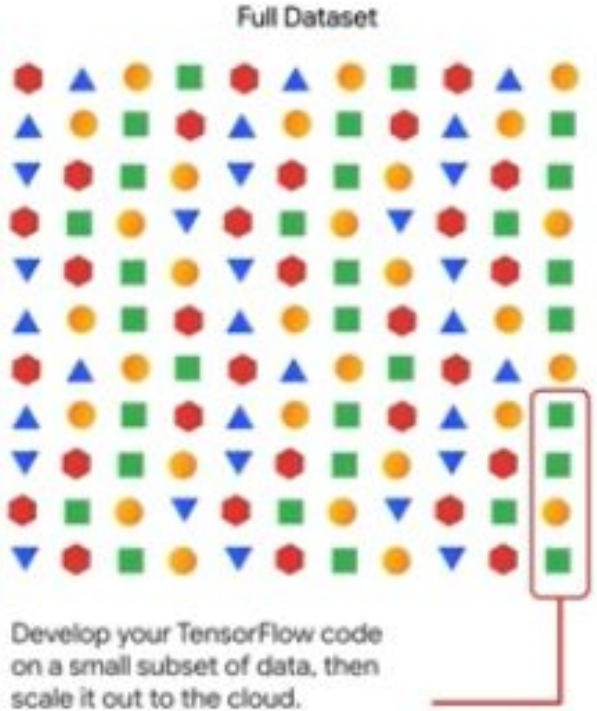
Use a confusion matrix to assess classification model performance



Evaluate the final model with independent test data



Developing the ML model software on the entire dataset can be expensive; you want to to develop on a smaller sample



Professional ML Engineer

Learning Journey

Google Developer Groups
Road to Google Developers Certification
Google Cloud

Pre-work	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Exam weeks
Exam guide review sessions ➪	kick-off session	exam guide review session	exam guide review session	exam guide review session	exam guide review session	exam guide review session + wrap up	**Self study (and potential exam)**
Review the Professional ML Engineer Exam Guide **Review the** Professional ML Engineer Sample Questions **Go through:** Google Cloud Platform Big Data and Machine Learning Fundamentals	**Complete course:** How Google does Machine Learning Launching into Machine Learning	**Complete course:** TensorFlow on Google Cloud Feature Engineering	**Complete course:** Machine Learning in the Enterprise	**Hands On Lab Practice:** Production Machine Learning Systems Computer Vision Fundamentals with Google Cloud	**Complete course:** Natural Language Processing on Google Cloud Recommendation Systems on GCP	**Complete course:** ML Ops - Getting Started ML Pipelines on Google Cloud **Check Readiness:** Professional ML Engineer Sample Questions **Hands On Lab Practice:** Perform Foundational Data, ML, and AI Tasks in Google Cloud (Skill Badge) - 7hrs Build and Deploy ML Solutions on Vertex AI (Skill Badge) - 8hrs	

Redeem your participation badge

Thank you for joining the event

Link to badge

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Thank you for tuning in!

For any operational questions about access to Cloud Skills Boost or the Road to Google Developers Certification program contact: gdg-support@google.com