AI IN ACTION

Building Your Essential AI Toolkit

The AI Landscape: Exploring the Latest Technologies

Joseph M. Reilly, PhD

Assistant Teaching Professor

College of Professional Studies

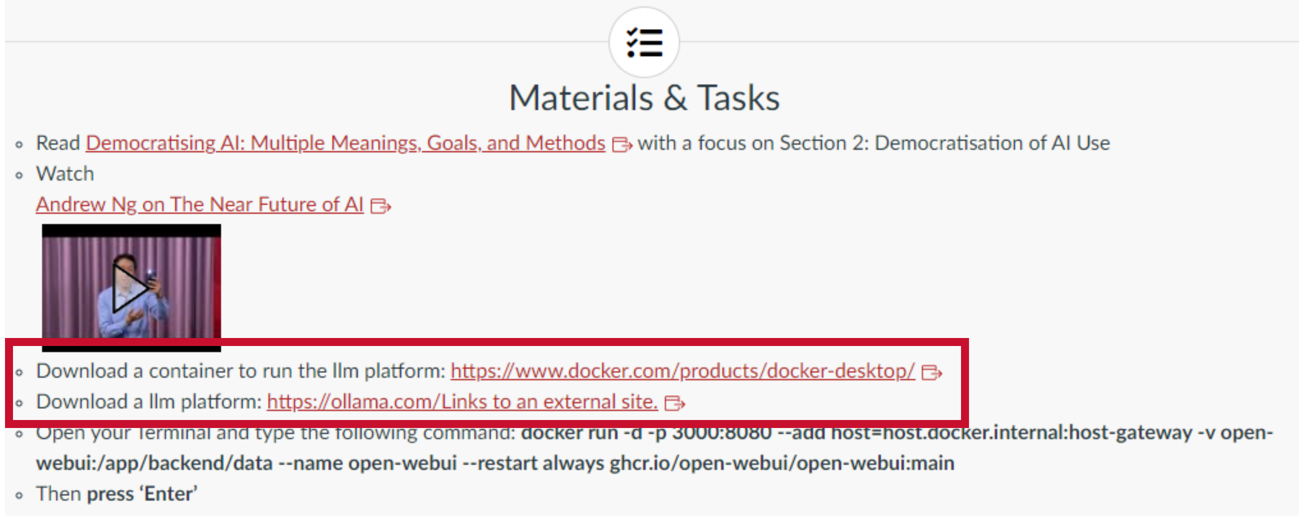
Northeastern University

Please download the demo materials!

Session 2: The AI Landscape: Exploring the Latest Technologies

↓

Materials & Tasks



Materials & Tasks

- Read Democratising AI: Multiple Meanings, Goals, and Methods with a focus on Section 2: Democratisation of AI Use
- Watch
  Andrew Ng on The Near Future of AI
- Download a container to run the llm platform: https://www.docker.com/products/docker-desktop/
- Download a llm platform: https://ollama.com/Links to an external site.
- Open your Terminal and type the following command: **docker run -d -p 3000:8080 --add host=host.docker.internal:host-gateway -v open-webui:/app/backend/data --name open-webui --restart always ghcr.io/open-webui/open-webui:main**
- Then **press 'Enter'**

Agenda for First Portion

1 The Current AI Market (5 mins)

2 Explainable AI and Open Source (10 mins)

3 Low-code and No-code AI (10 mins)

4 Example: Google Vertex AI (5 mins)

The Current AI Market

Where we're at and where we're going (maybe)

AI Market Expansion

2024

$638b

2034

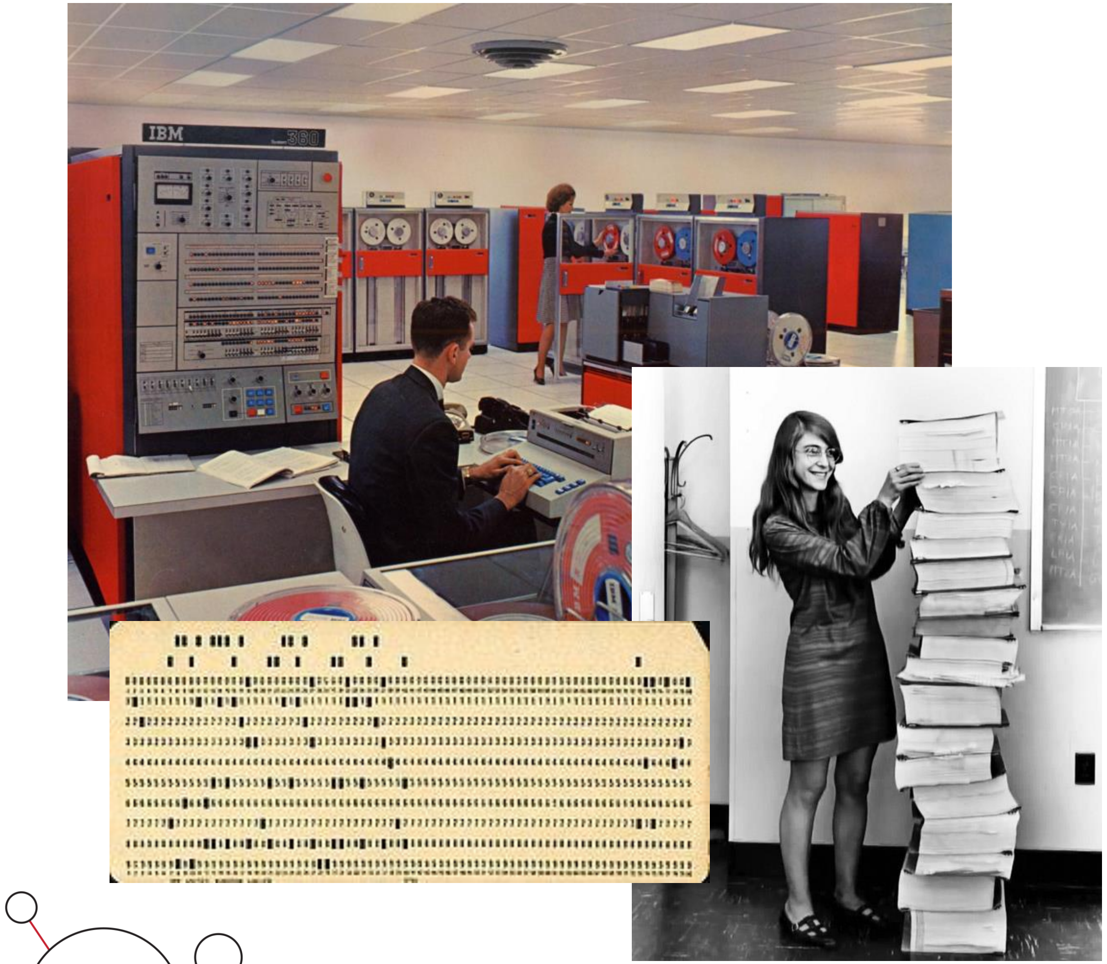
$3,680b

- North America largest market share, but other regions catching up
- Heavy investment by tech giants, demand from end-use verticals
- Favorable government initiatives and little regulation
- Lack of skilled professionals to leverage AI
  - Augmenting the workforce, not automating it away

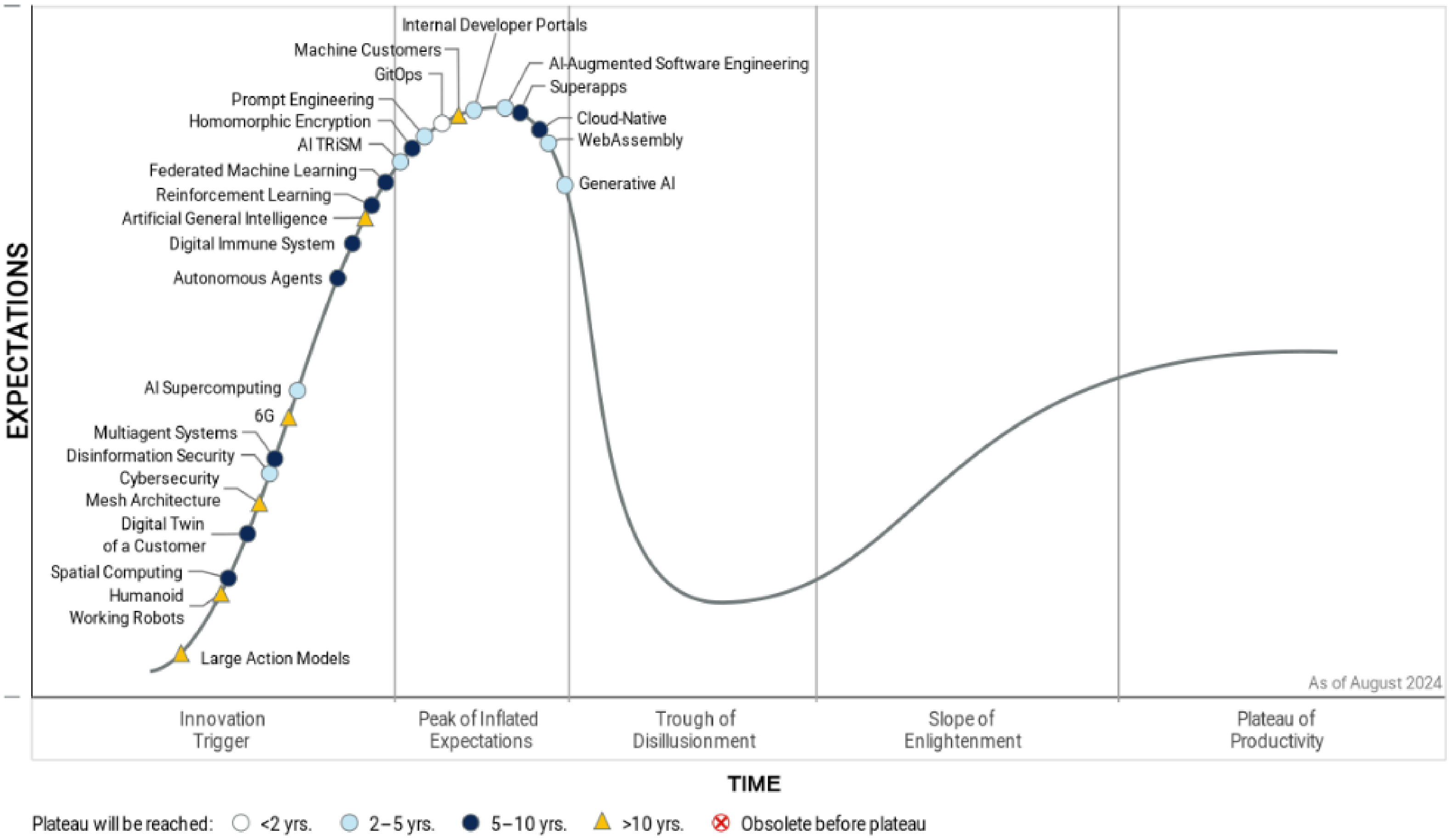
AI IN ACTION
Building Your Essential AI Toolkit

Precedence Research | Artificial Intelligence (AI) Market
Statista | Artificial Intelligence - Worldwide

Growth Driven by Increased Access



The Trough of Disillusionment



Gartner

Gartner 2024 Hype Cycle Press Release

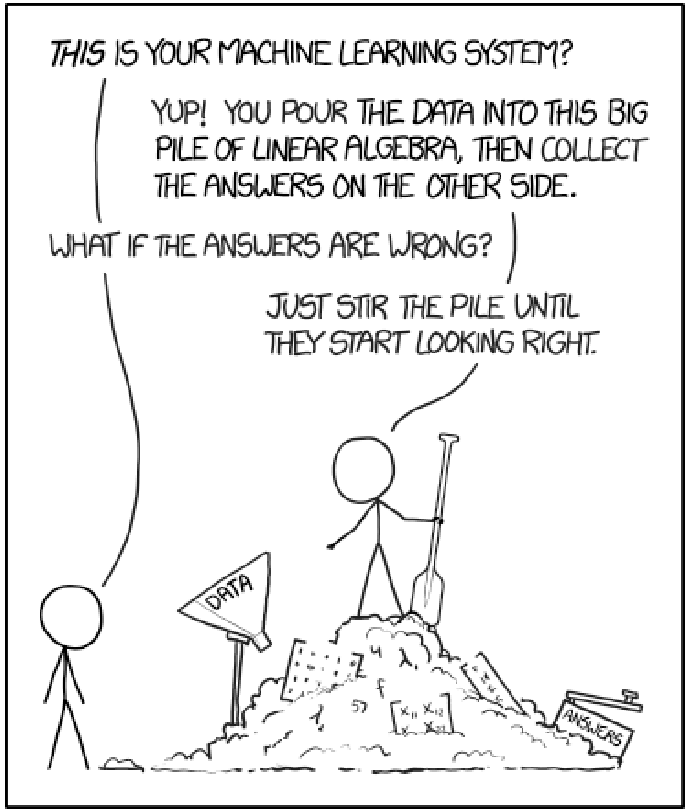
Explainable AI and Open Source

Explain the behavior and see how we got there

The Danger of the Black Box

Data → Prediction

- If you can’t explain why a model arrived at a specific decision, how can you defend it?
- Big issue in deep learning
- How can you trust it?
  - Mortgage lending
  - Medical diagnoses
  - Predictive policing
  - Hiring and firing


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

xkcd | Machine Learning
HDSR | Why Are We Using Black Box Models in AI When We Don’t Need To?

Explainable AI (XAI)

Interpretability

- Observe the inner workings
- Cause and effect of each parameter and input
- Very difficult for complex, high performance models

Explainability

- Explain the behavior in human terms beyond weights and features
- Indirect, model-agnostic ways to find meaning from results
- Communicate to end users

Must accuracy be sacrificed for interpretability?

Not necessarily.

AI IN ACTION
Building Your Essential AI Toolkit

AWS | Interpretability versus explainability
IBM | What is explainable AI?

Open-Source AI

Harder to define than you'd think!

Current draft from the Open Source Initiative

Free software means the freedom to…

- run the program as you wish
- study how the program works and change it as you wish
- redistribute copies to help others
- distribute copies of your modified versions to others

Complications with AI:

- Free access to the training data?
- Use-based licenses?
- Code related to preprocessing?

GNU | What is free software?
MIT Tech Review | The tech industry can't agree on what open-source…

Benefits of Open Source

Transparency

Understand how it works and mitigate biases

Customization

Fine-tune models on sensitive data

Collaboration

Understand how it works and mitigate biases

Lower Barriers to Use

Efficiency and affordability to experiment and deploy

Meta | Open-Source AI is the Path Forward
Google Cloud | Open-source AI

Open-Source AI Examples

	Open Source	Closed Source
Enterprise	NVIDIA Meta	OpenAI (Microsoft) Google (Alphabet) amazon
Smaller	mosaicML together.ai Mistral AI stability.ai Hugging Face	cohere AI21labs ADEPT Inflection ANTHROP\C

[Towards AI | The Choice for Businesses between Open-Source and Proprietary Models...]

Low-code and No-code AI

Democratizing AI and empowering new users

Emergence of Low-Code Options

- Visual approach to development to reduce traditional programming
  - Traces back to late 80's / early 90's with Visual Basic and web development
- Helpful for experienced devs as well as novices
- Connecting with existing ML/AI tools

20% growth
from '22 to '23

80% of user base
outside of IT dept

70% of new apps
will use low code by '25

Gartner | Gartner Forecasts Wordwide Low-Code Development Technology…
IBM | What is low-code?

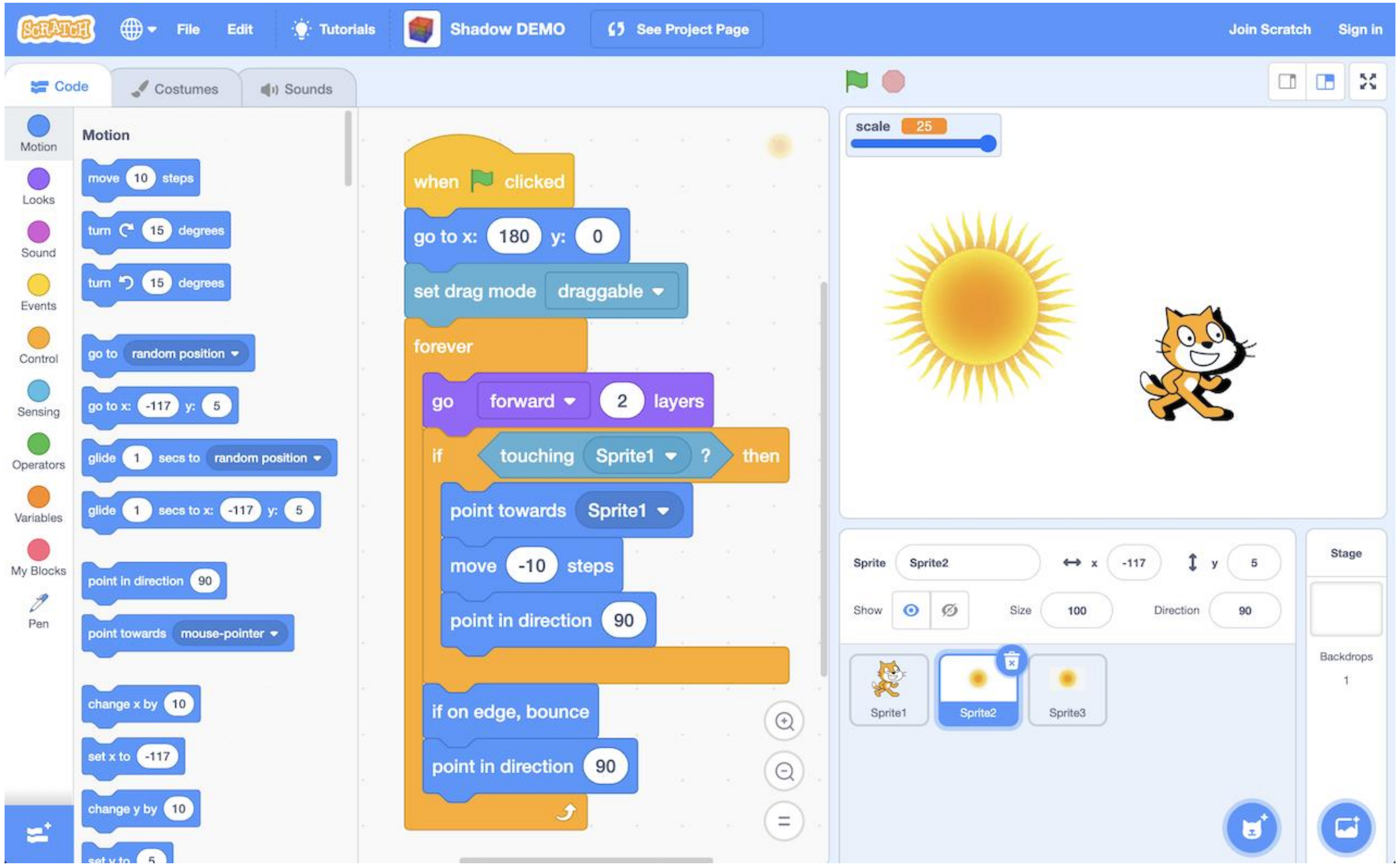
Advantages of Enterprise Low-Code

- Agility
- Less Tech Debt
- Lower Costs
- Speeds Up Dev
- Citizen Developer
- Levels of Abstraction

AI IN ACTION
Building Your Essential AI Toolkit

Bruhin et al. | The Rise of Generative AI in Low Code Development Platforms
Forbes | 14 Benefits Of Low-Code App Development That Tech Pros Love

Scratch!



```
when 🏴 clicked
go to x: 180 y: 0
set drag mode draggable
forever
    go forward 2 layers
    if touching Sprite1 ? then
        point towards Sprite1
        move -10 steps
        point in direction 90
    if on edge, bounce
    point in direction 90
```

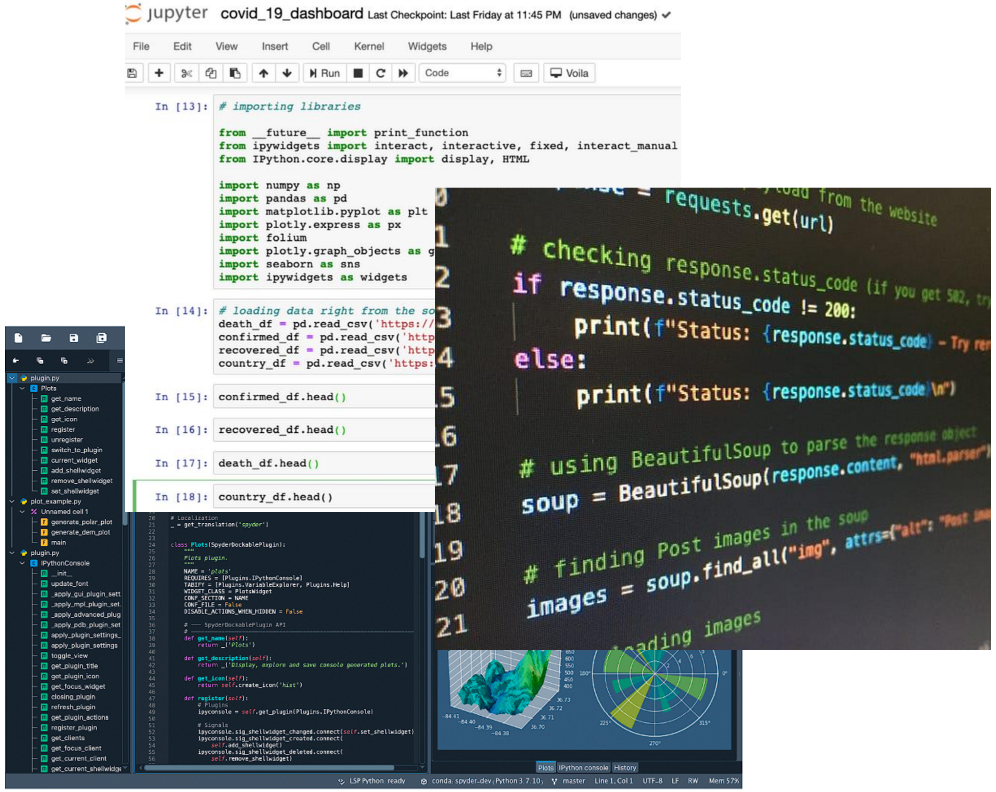
scale 25

Sprite Sprite2 ↔ x -117 ↕ y 5

Show Size 100 Direction 90

scratch.mit.edu

Evolution of AI Development

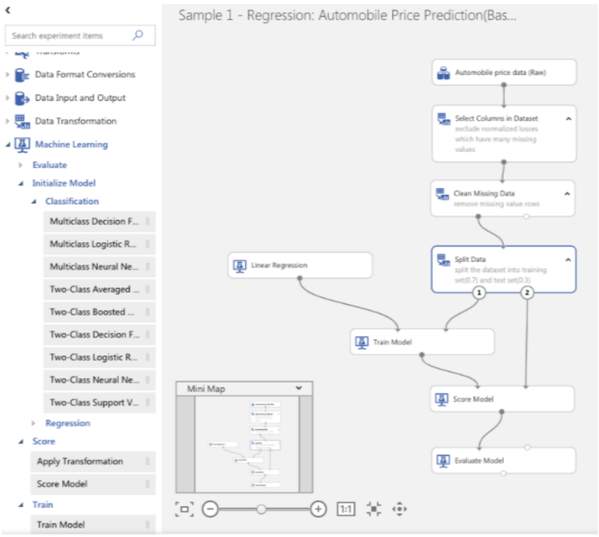


covid_19_dashboard Last Checkpoint: Last Friday at 11:45 PM (unsaved changes)

```
In [13]: # importing libraries

         from __future__ import print_function
         from ipywidgets import interact, interactive, fixed, interact_manual
         from IPython.core.display import display, HTML

         import numpy as np
         import pandas as pd
         import matplotlib.pyplot as plt
         import plotly.express as px
         import folium
         import plotly.graph_objects as g
         import seaborn as sns
         import ipywidgets as widgets

In [14]: # loading data right from the so
         death_df = pd.read_csv('https://
         confirmed_df = pd.read_csv('http
         recovered_df = pd.read_csv('http
         country_df = pd.read_csv('https:

In [15]: confirmed_df.head()

In [16]: recovered_df.head()

In [17]: death_df.head()

In [18]: country_df.head()
```

```
                  requests.get(url)

# checking response.status_code (if you get 502, try
if response.status_code != 200:
    print(f"Status: {response.status_code} - Try re
else:
    print(f"Status: {response.status_code}\n")

# using BeautifulSoup to parse the response object
soup = BeautifulSoup(response.content, "html.parser")

# finding Post images in the soup
images = soup.find_all("img", attrs={"alt": "Post im

  loading images
```

```
# Localization
_ = get_translation('spyder')


class Plots(SpyderDockablePlugin):
    """
    Plots plugin.
    """
    NAME = 'plots'
    REQUIRES = [Plugins.IPythonConsole]
    TABIFY = [Plugins.VariableExplorer, Plugins.Help]
    WIDGET_CLASS = PlotsWidget
    CONF_SECTION = NAME
    CONF_FILE = False
    DISABLE_ACTIONS_WHEN_HIDDEN = False

    # --- SpyderDockablePlugin API
    # ------------------------------------------------------------
    def get_name(self):
        return _('Plots')

    def get_description(self):
        return _('Display, explore and save console generated plots.')

    def get_icon(self):
        return self.create_icon('hist')

    def register(self):
        # Plugins
        ipyconsole = self.get_plugin(Plugins.IPythonConsole)

        # Signals
        ipyconsole.sig_shellwidget_changed.connect(self.set_shellwidget)
        ipyconsole.sig_shellwidget_created.connect(
            self.add_shellwidget)
        ipyconsole.sig_shellwidget_deleted.connect(
            self.remove_shellwidget)
```

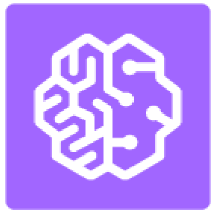


Sample 1 - Regression: Automobile Price Prediction(Bas...

Automobile price data (Raw)
Select Columns in Dataset
exclude normalized losses which have many missing values
Clean Missing Data
remove missing value rows
Linear Regression
Split Data
split the dataset into training set(0.7) and test set(0.3)
1
2
Train Model
Score Model
Evaluate Model

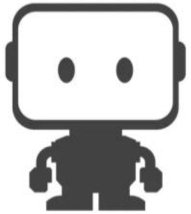
Tune hyperparameters of a Random Forest Classifier on my training data, then report the precision and recall on the test set. Explain the results concisely.

Representative Sample of Some Players

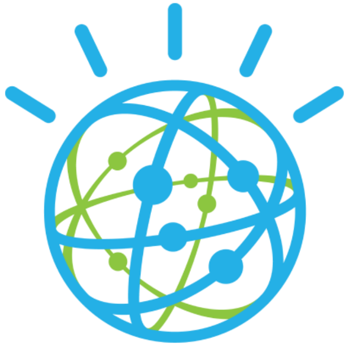


Amazon SageMaker

Azure Machine Learning



DataRobot



IBM Watson

Example: Google Vertex AI

See how this works in practice at enterprise scale

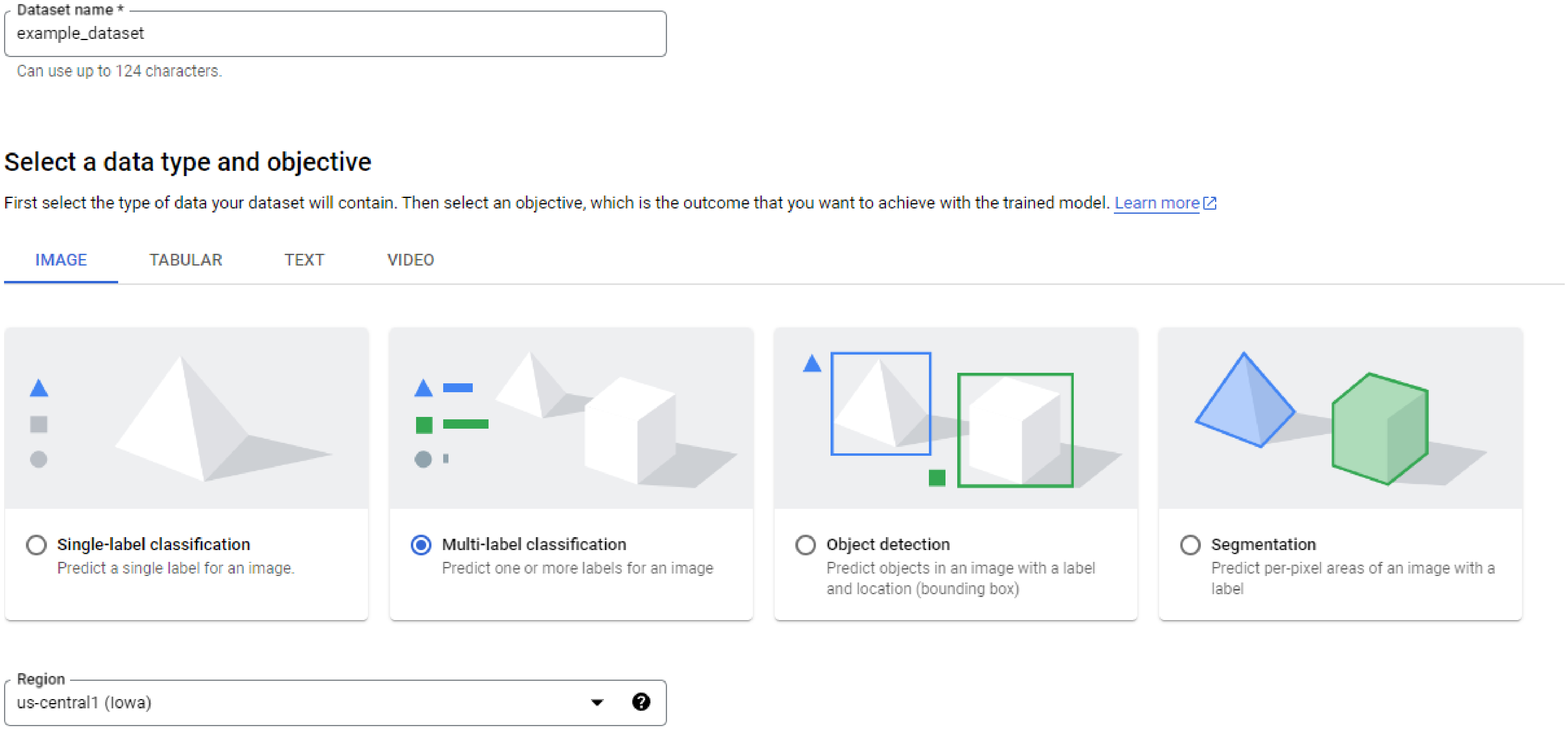
Important Caveats!

I do not work for Google.

They aren’t paying me to present this.

I don’t currently use Vertex AI but I have used GCP.

It all starts with data



Dataset name *
example_dataset
Can use up to 124 characters.

Select a data type and objective

First select the type of data your dataset will contain. Then select an objective, which is the outcome that you want to achieve with the trained model. Learn more

○ **Single-label classification**
Predict a single label for an image.

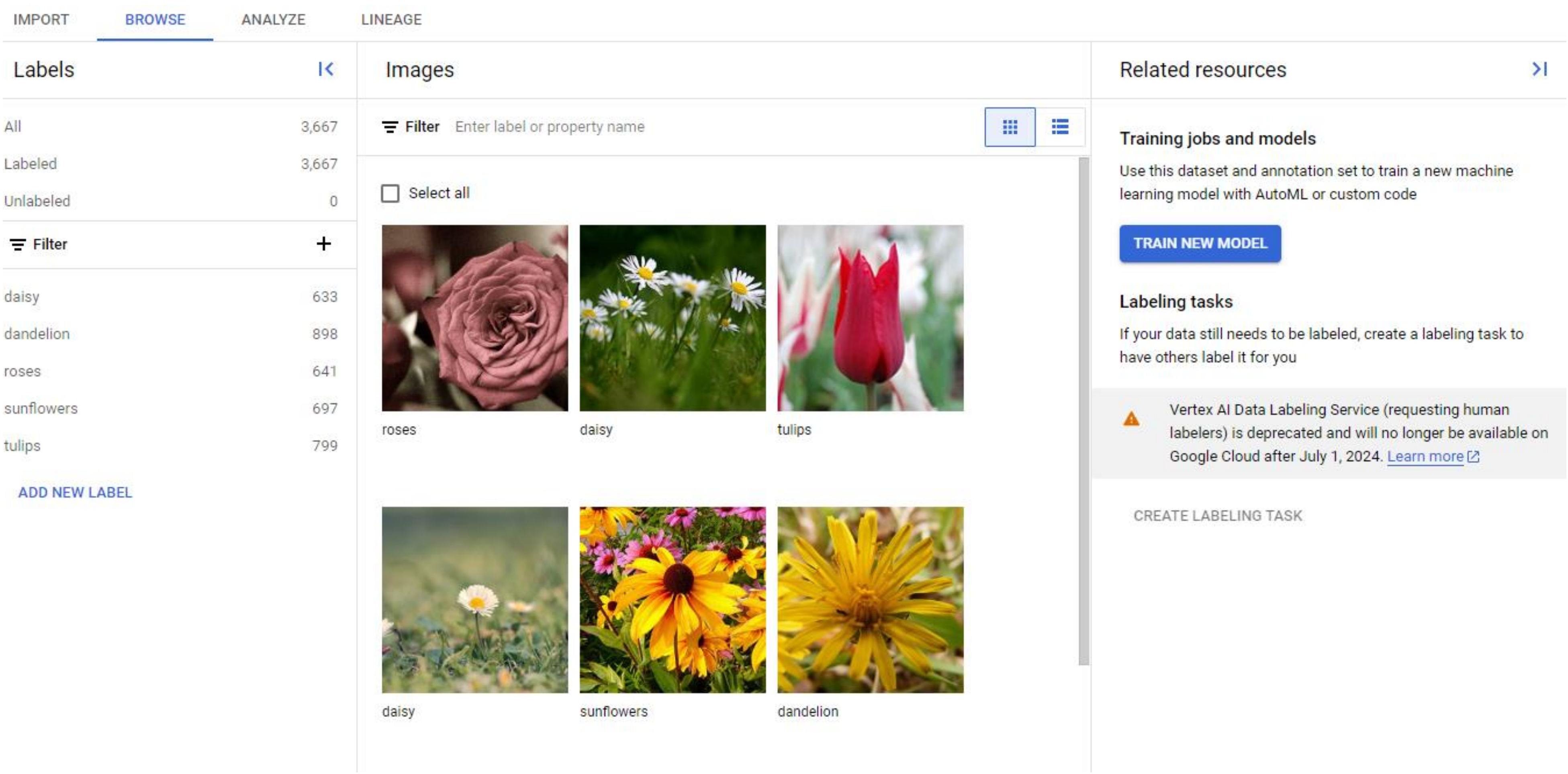
● **Multi-label classification**
Predict one or more labels for an image

○ **Object detection**
Predict objects in an image with a label and location (bounding box)

○ **Segmentation**
Predict per-pixel areas of an image with a label

Region
us-central1 (Iowa)

It all starts with data



Labels

All	3,667
Labeled	3,667
Unlabeled	0

Filter +

daisy	633
dandelion	898
roses	641
sunflowers	697
tulips	799

ADD NEW LABEL

Images

Filter Enter label or property name

□ Select all

roses daisy tulips

daisy sunflowers dandelion

Related resources

Training jobs and models

Use this dataset and annotation set to train a new machine learning model with AutoML or custom code

TRAIN NEW MODEL

Labeling tasks

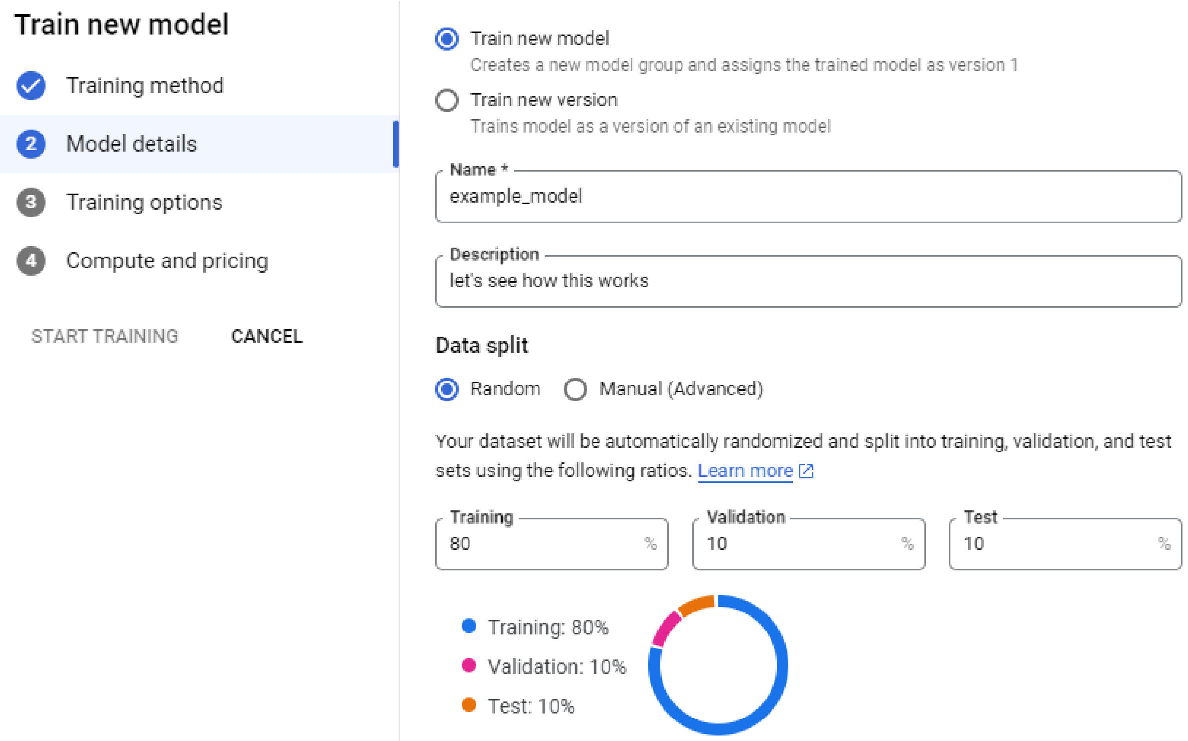
If your data still needs to be labeled, create a labeling task to have others label it for you

⚠ Vertex AI Data Labeling Service (requesting human labelers) is deprecated and will no longer be available on Google Cloud after July 1, 2024. Learn more

CREATE LABELING TASK

AI IN ACTION
Building Your Essential AI Toolkit

Model training on the cloud



Train new model

- ✓ Training method
- 2 Model details
- 3 Training options
- 4 Compute and pricing

START TRAINING CANCEL

● Train new model
Creates a new model group and assigns the trained model as version 1

○ Train new version
Trains model as a version of an existing model

Name *: example_model

Description: let's see how this works

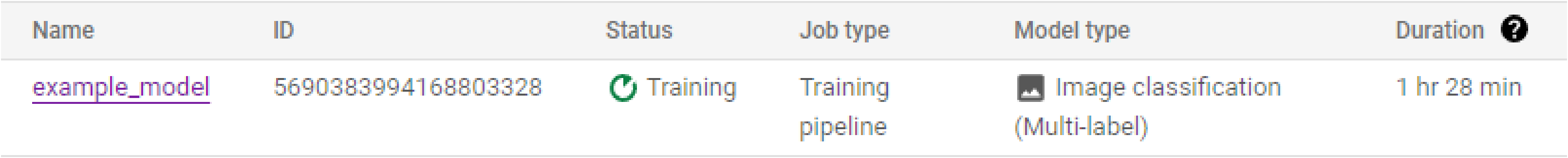
Data split

● Random ○ Manual (Advanced)

Your dataset will be automatically randomized and split into training, validation, and test sets using the following ratios. Learn more

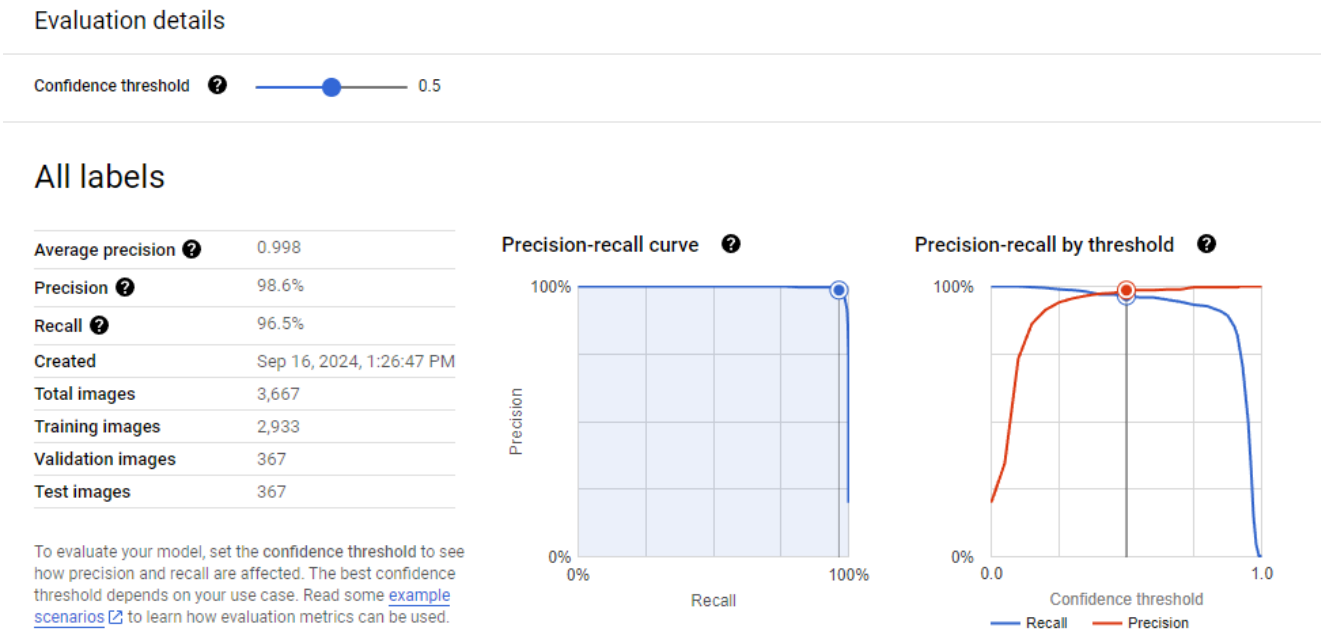
Training: 80 %
Validation: 10 %
Test: 10 %

- ● Training: 80%
- ● Validation: 10%
- ● Test: 10%



Name	ID	Status	Job type	Model type	Duration
example_model	5690383994168803328	Training	Training pipeline	Image classification (Multi-label)	1 hr 28 min

Model evaluation



Evaluation details

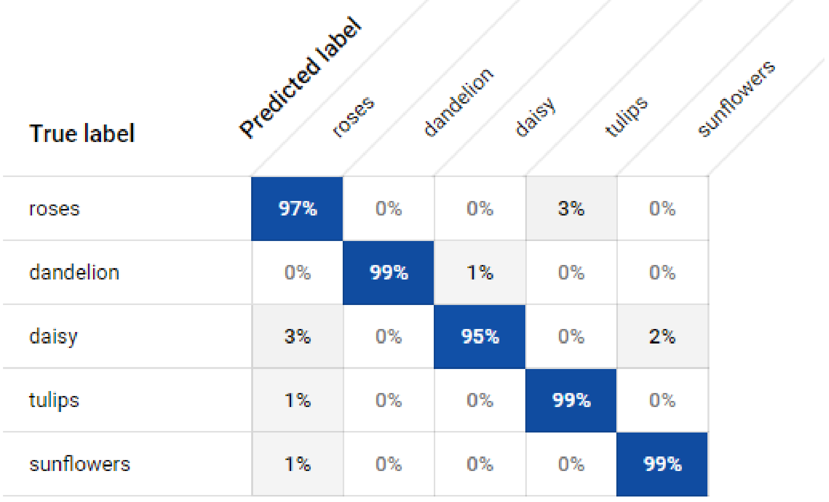
Confidence threshold 0.5

All labels

Average precision	0.998
Precision	98.6%
Recall	96.5%
Created	Sep 16, 2024, 1:26:47 PM
Total images	3,667
Training images	2,933
Validation images	367
Test images	367

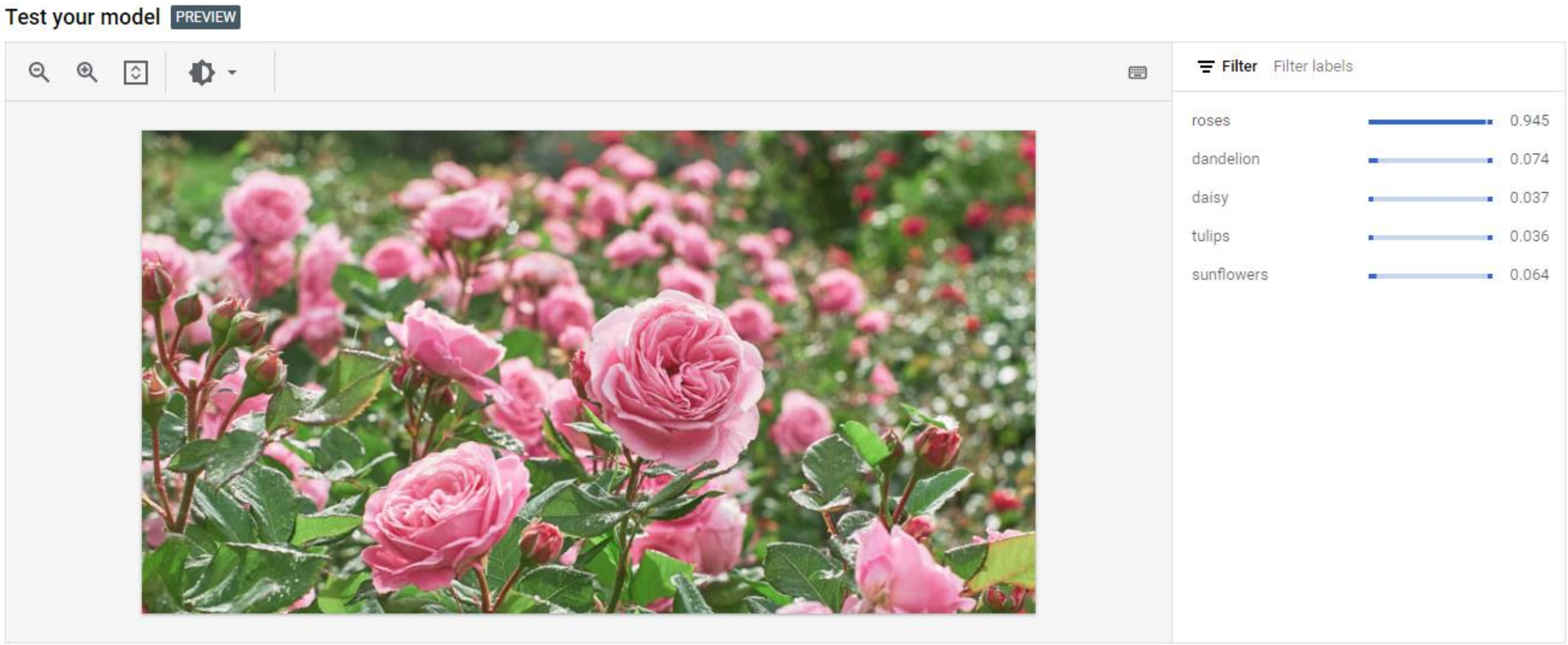
To evaluate your model, set the **confidence threshold** to see how precision and recall are affected. The best confidence threshold depends on your use case. Read some example scenarios to learn how evaluation metrics can be used.

Precision-recall curve
100%
0%
Precision
0%
100%
Recall
Precision-recall by threshold
100%
0%
0.0
1.0
Confidence threshold
Recall
Precision



True label	Predicted label: roses	dandelion	daisy	tulips	sunflowers
roses	97%	0%	0%	3%	0%
dandelion	0%	99%	1%	0%	0%
daisy	3%	0%	95%	0%	2%
tulips	1%	0%	0%	99%	0%
sunflowers	1%	0%	0%	0%	99%

Model deployment

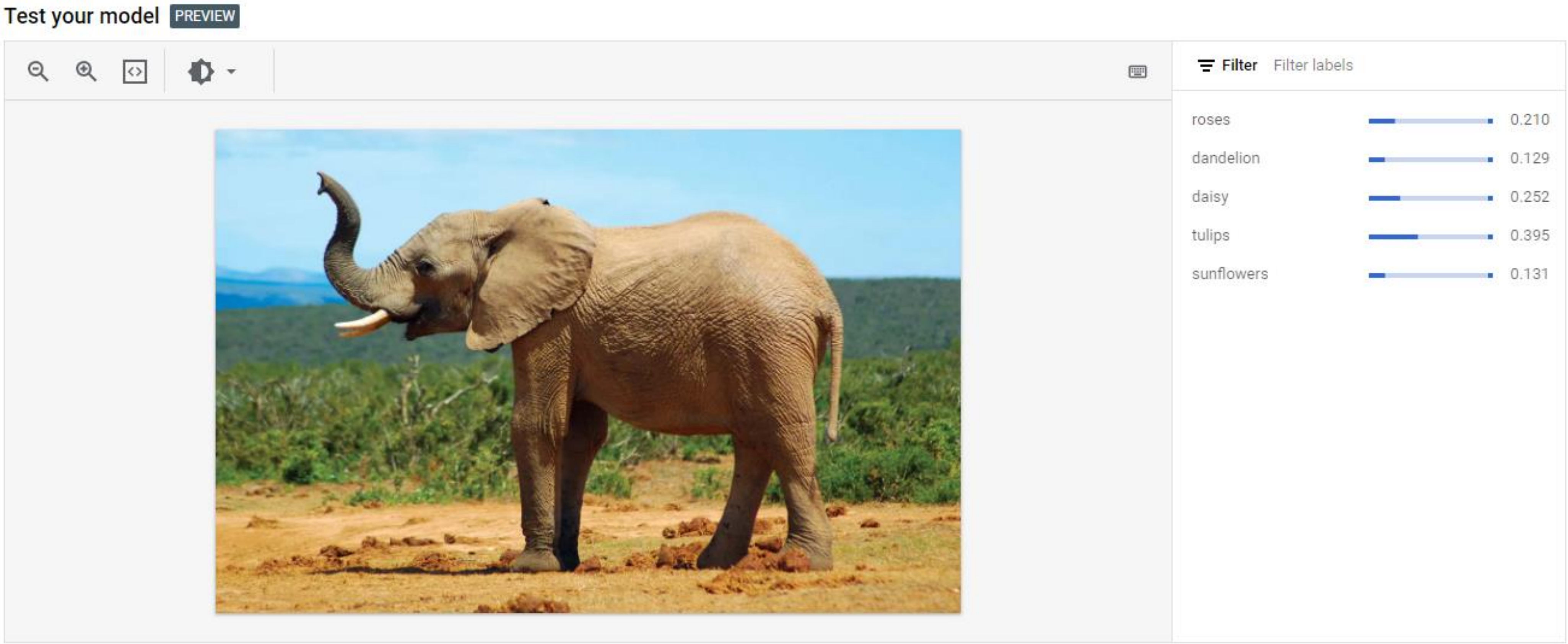


Test your model PREVIEW

Filter Filter labels

roses	0.945
dandelion	0.074
daisy	0.037
tulips	0.036
sunflowers	0.064

Model deployment



Test your model PREVIEW

Filter Filter labels

roses	0.210
dandelion	0.129
daisy	0.252
tulips	0.395
sunflowers	0.131

AI IN ACTION

Building Your Essential AI Toolkit

Thank you!

Northeastern University