



Machine Learning Decisions and Deployment

Data Boot Camp

Lesson 21.3



Class Objectives

By the end of this lesson, you will be able to:



Assess the trade-offs between machine learning models.



Choose and build an appropriate machine learning model for a given dataset and business case.



Design an appropriate machine learning pipeline.



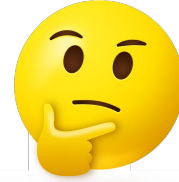
Create and deploy a machine learning pipeline.

Revisiting the Great Debate

Thought Experiment #1

Instructor Do: Revisiting the Great Debate

Which do Americans prefer: Italian or Mexican food?





Which steps of the analytics paradigm could benefit from machine learning?

Revisiting the Great Debate

Here are some ideas:

01

Decompose the question.

02

Identify data sources.

03

We could build a pipeline to automatically retrieve and clean the data.

04

Analyse for trends.

05

Acknowledge limitations of the analysis / make the call or tell the story.



Groups Do: The Great Debate Revisited

In groups, come up with another strategy to try and answer the question: **“Which do Australians prefer: Italian or Thai food?”**
Consider what trade-offs need to be made, and be ready to discuss with the class.

Suggested Time:

15 minutes



Time's Up! **Let's Review.**

Questions?



A Better Question for Machine Learning



We'll tweak the question from the previous activity to focus on something more appropriate for machine learning.

Thought Experiment 2

You are interviewing at **yelp** 

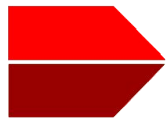
As part of the process, they have asked you to complete a take-home assignment:

Create and train a model to predict a Yelp user's average rating of Italian restaurants.



Thought Experiment 2

Assume the following:



The model must use data from the [Yelp dataset](#).



You will have one week to work on the assignment.



You may use external libraries, but you must be able to explain how they work at a high level.



You should be prepared to explain how your model works to interviewers of varying technical backgrounds.

We'll base our analysis
on a real dataset, but this dataset
is over 6 GB compressed.



Instead of writing any code or
downloading the data, we're going to make a
hypothetical plan and rely on the data description
from Yelp to guide how we will approach the data.



Groups Do: Preference Predictor

Break into groups again, and develop a plan to create a machine learning model that predicts a Yelp user's average rating of Italian restaurants. Be ready to discuss your ideas with the class.

Suggested Time:

25 minutes



Time's Up! **Let's Review.**

Questions?





Time to Code

Preference Predictor Critique

Suggested Time:

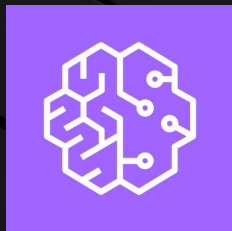
15 Minutes

Questions?



A close-up, high-angle shot of a computer keyboard. The central focus is a large, white, rectangular key with rounded corners. On this key, there is a dark blue icon of a coffee cup with three wavy lines above it representing steam. Below the icon, the word "Break" is printed in a dark blue, serif font. The key is set against a background of other keyboard keys, which are slightly out of focus. The lighting is soft and even, highlighting the texture of the keys and the wood-grain surface of the keyboard base.

Break



Amazon SageMaker Introduction and Setup

Amazon SageMaker Introduction and Setup



Up until now, all the models we've created run within a Jupyter notebook.

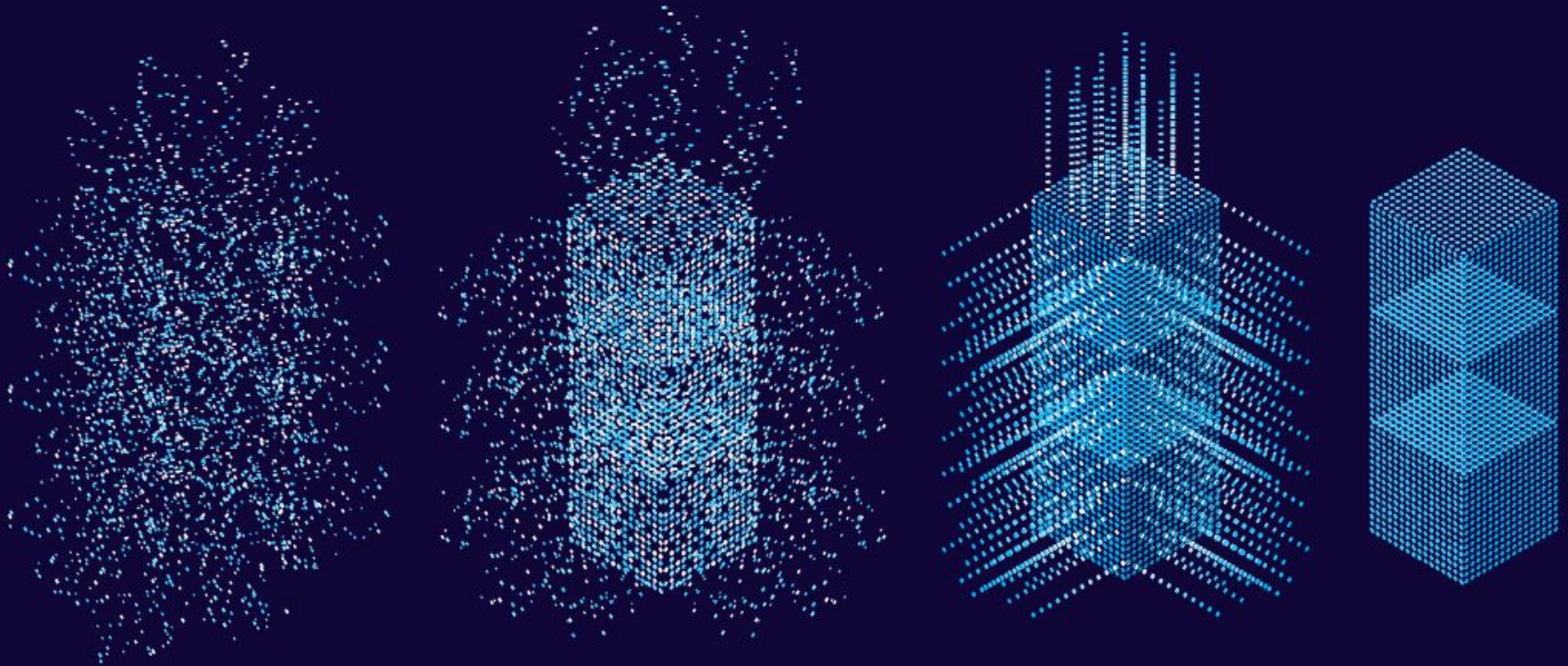


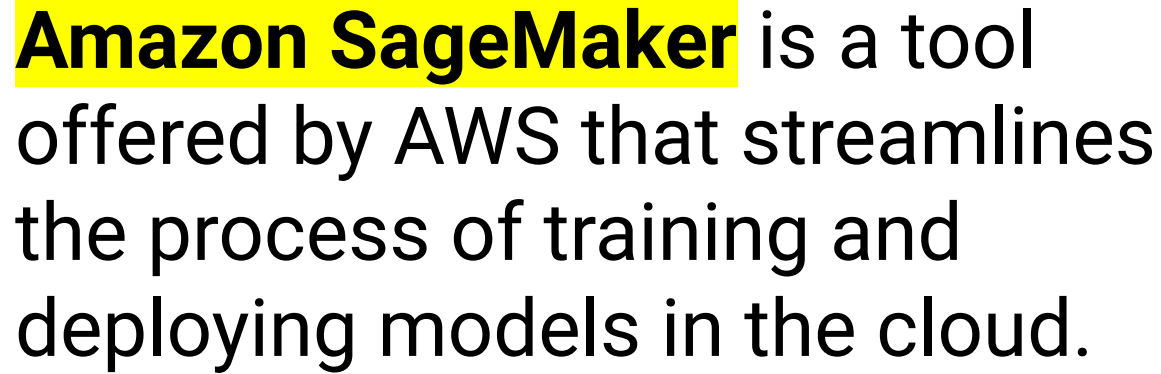
If we want to share the functionality of our models with other people (or other programs), we need to deploy them.



One way to deploy models is to build an API from scratch using Flask, but that takes a lot of work, plus the API needs to be hosted on a server.

The rest of the session will be about deploying machine learning models using **Amazon SageMaker**.





Amazon SageMaker is a tool offered by AWS that streamlines the process of training and deploying models in the cloud.



Instructor Demonstration

Creating an Administrator User on IAM



A standard best practice is to avoid using the principal, or *root*, user to manage your AWS account.



Pro Tip:

Use the AWS Identity and Access Management (IAM) service to create a new user for each person that requires administrator access.

Questions?





Activity: Creating an Admin User on IAM

In this activity, you will create an administrator user to manage your AWS account.

Suggested Time:

10 minutes



Time's Up! Let's Review.

Amazon SageMaker Notebooks



Time to Code

Create an Amazon SageMaker Notebook Instance

Suggested Time:

20 Minutes



The first component that
Amazon SageMaker requires
is an Amazon S3 bucket.

Storage with AWS S3

S3 uses **buckets** to store files, which are similar to computer folders or directories. Buckets can contain additional folders and files. Each bucket must have a unique name.

S3 has fine-grained control over files, such as read and write permissions. Buckets can assign individual access or total public access.



The S3 bucket structure is somewhat similar to a GitHub repository, which also holds files and folders.



Storage with AWS S3

An S3 bucket can contain files, but it cannot contain another bucket.



Questions?





Time to Code

Create and Deploy a Machine Learning Model in Amazon SageMaker

Suggested Time:

30 Minutes

Questions?



Pros and Cons of Deploying Machine Learning Models with Amazon SageMaker



Pros of Deploying Machine Learning Models with Amazon SageMaker

Data storage capacity:	By using an Amazon S3 bucket to store the data...	We could have trained a model on multiple terabytes of data, or a lot more space than would otherwise have fit in our personal computer.
Hardware/GPU:	By using different Amazon SageMaker instances to train our model...	We can access compute power, including GPU capabilities, making powerful hardware available to us as required.
Cost:	Using AWS resources...	We only pay for what we use. We'll turn off everything before ending the class and not incur further charges.
Availability:	By deploying our model to another Amazon SageMaker instance...	We have made the prediction functionality available 24/7 through a secured endpoint to an application, or which can be consumed by others without having to make our computer available.
RESTful API:	As we discussed in previous units...	APIs provide a standard mechanism to access data. Our ML API can be consumed through apps and other channels in a simple form while remaining secure and allowing other constraints (e.g., authentication, authorization, rate limiting, etc.).



Cons of Deploying Machine Learning Models with Amazon SageMaker

Data privacy/ security

By uploading data to a third party, you are trusting your data to them. Certain kinds of data are subject to compliance and regulatory constraints.

Visibility:

You won't have oversight on AWS internal handling of your data and infrastructure.

Availability:

Although there are SLAs in place, AWS (and any other cloud providers) can and have suffered outages at times, causing data unavailability.

Questions?



SageMaker Clean-Up



Instructor Demonstration

Delete Notebook Instance

Questions?





Time to Code



Delete AWS Resources

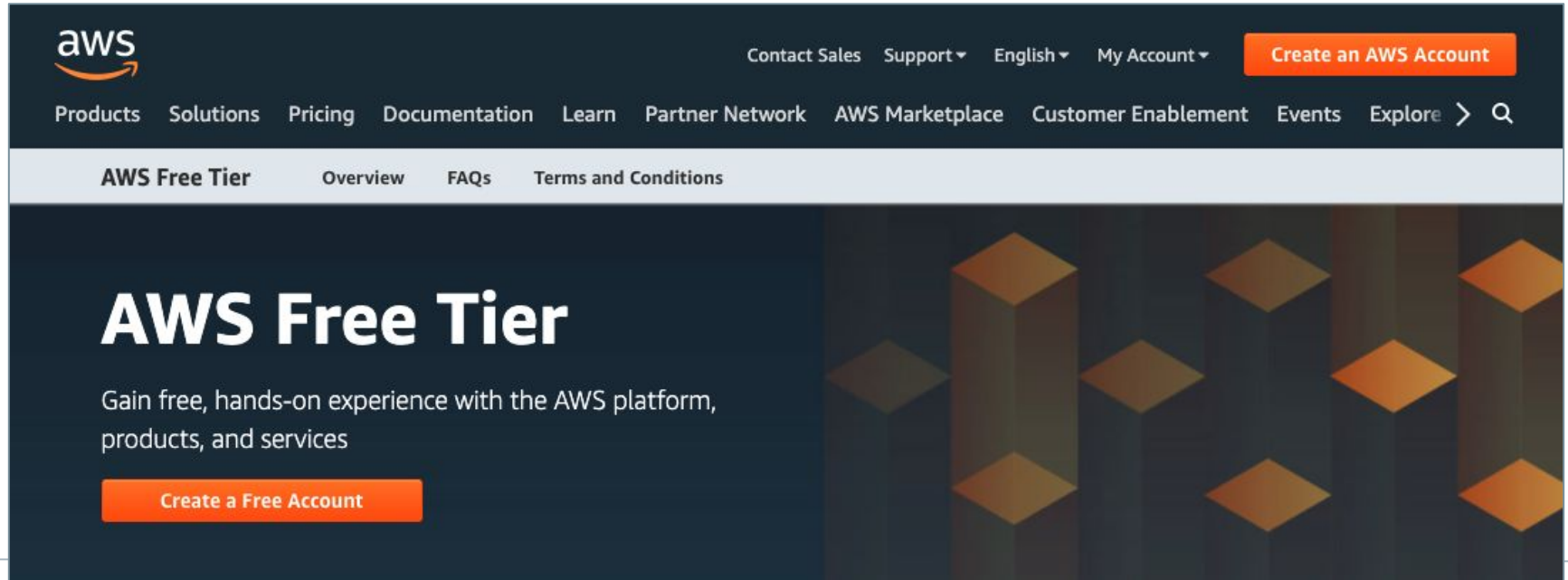
Suggested Time:

10 Minutes

Delete AWS Resources

Policies for the AWS free tier and trials continually change.

You must remove any unnecessary resources created on AWS to avoid additional charges, especially Amazon SageMaker instances because AWS bills you for hosting the instances (even if they are stopped).

A screenshot of the AWS Free Tier landing page. The page has a dark blue header with the AWS logo on the left and navigation links (Contact Sales, Support, English, My Account) and a 'Create an AWS Account' button on the right. Below the header is a horizontal menu with links: Products, Solutions, Pricing, Documentation, Learn, Partner Network, AWS Marketplace, Customer Enablement, Events, Explore, and a search icon. A light blue sub-header contains 'AWS Free Tier', 'Overview', 'FAQs', and 'Terms and Conditions'. The main content area features the title 'AWS Free Tier' in large white font, followed by the text 'Gain free, hands-on experience with the AWS platform, products, and services'. At the bottom left is a 'Create a Free Account' button. The background of the main content area is dark blue with a pattern of 3D isometric cubes in various shades of blue and orange.

aws

Contact Sales Support English My Account Create an AWS Account

Products Solutions Pricing Documentation Learn Partner Network AWS Marketplace Customer Enablement Events Explore > Q

AWS Free Tier Overview FAQs Terms and Conditions

AWS Free Tier

Gain free, hands-on experience with the AWS platform, products, and services

Create a Free Account

Delete AWS Resources

Follow these [clean-up steps](#) on AWS to delete AWS resources. Be sure to remove ALL of the following:



Endpoint



Endpoint configuration



Model

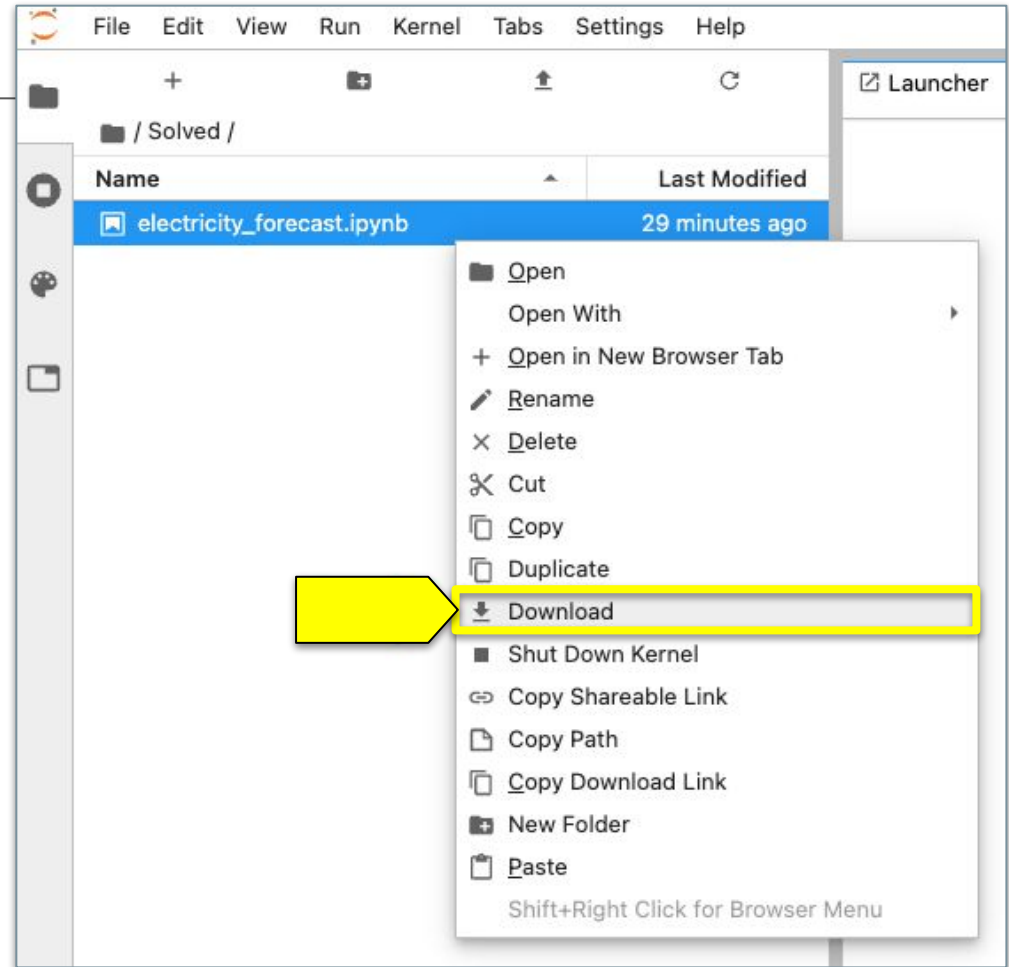


Notebook instance

Delete AWS Resources

Remember:

you can save a local copy of the Jupyter notebooks by right-clicking on the notebook name and selecting the Download option.



Questions?



*The
End*